İstanbul Üniversitesi
Istanbul University

ISSN 1303-0914
Vol. 17, No. 1, 2017

Istanbul University-Journal of
ELECTRICAL & ELECTRONICS ENGINEERING
(IU-JEEE)

Publication of Istanbul University Engineering Faculty

Editorial to the Special Issue on Selected Papers from Medical Technologies Conference, TIPTEKNO-2016

This special issue of the *Istanbul University Journal of Electrical and Electronics Engineering, IU-JEEE* contains 14 selected papers from 2016 Medical Technologies Conference, TIPTEKNO-2016 which was held on October 27-29, 2016 in Belek, Antalya, Turkey with international attendance. Furthermore, 14 regular papers are also included to content of this issue.

TIPTEKNO conferences have been organized since 2010 by the Turkish Biomedical and Clinical Engineering Association. TIPTEKNO-2016 was technically co-sponsored by IEEE Turkey Section, Engineering in Medicine and Biology Society (EMBS) Chapter.

TIPTEKNO Conferences aim to bring together medical technology users, developers, researchers, academicians, managers, and government officers. In addition to cutting edge research paper presentations in biomedical and clinical engineering areas, the conference serves as a multi-disciplinary platform to discuss the problems in healthcare industry, and clinical engineering regulations. Every year the attendees of the conference are exposed to panel discussions on important technical and industrial problems, training seminars on recent scientific developments, as well as lecture and poster presentations on research papers.

180 papers have been submitted for presentation at TIPTEKNO-2016. Based on review reports by experts in the Technical Program Committee, 110 papers have been accepted for presentation at the conference. The topics of the presented papers covers a wide range of fields in biomedical engineering; from signal and image processing to biological system modelling, from medical device development to bio- nano- materials, and in clinical engineering from hospital management to medical calibration systems.

30 papers from the above studies have been selected based on the review results, and invited to submit their extended versions to this special issue. After the Journal’s standard review process, we finalize this special issue covering the latest research results in Biomedical and Clinical Engineering.

We would like to thank the contributing authors for accepting our invitation, and also to the anonymous reviewers who helped us ensure the high technical quality of the accepted papers.

I hope you enjoy reading this *TIPTEKNO-2016 Special Issue of Istanbul University Journal of Electrical and Electronics Engineering, IU-JEEE.*

Aydin Akan, Editor

CONTENTS

CONTENTS

Fabrication of Ultra-micro Carbon Fiber Electrode Probes for Detection of O_2 and H_2O_2

Mustafa ŞEN

Department of Biomedical Engineering, Izmir Katip Celebi University, Izmir, Turkey
mustafa.sen@ikc.edu.tr

Abstract: Carbon fiber electrodes (CFEs) are commonly used in detection of neurotransmitters like dopamine. Besides, modification of these electrodes with enzymes enables development of biosensors capable of local analysis. Here, CFEs were fabricated using glass capillary tubes. Basically carbon fibers were inserted into glass capillary tubes and then the tubes were pulled using a micro-puller to insulate carbon fibers. Subsequently, the electrode surface was modified with Pt nanoparticles to evaluate the potential of these electrodes in detection of O_2 and H_2O_2. In the future, these electrodes will be used for construction of biosensors for detection of local ATP.

Keywords: *Carbon fiber electrode, H_2O_2, biosensor, electrochemical analysis.*

1. Introduction

In general, biosensors are analytical devices with a biological detection unit capable of sensing physiologically important molecules with the desired precision and converting them into numerical data. Enzymes, antibodies, microorganisms, cells, nucleic acids and sometimes tissues are used as biological detection units in biosensors [1-4]. These types of devices are designed mostly in the form of microchips or probes, while being completely unique to the investigator. In recent years, research has concentrated on making biosensors more efficient and cheaper, with higher sensitivity. The method used in the detection may be optical, mechanical or electrochemical. Although most of the biosensors developed so far are optically based, there are some disadvantages to this technique, such as fluctuations due to quenching or emission from non-target molecules, the shielding effect by the turbid solution and the need to mark the target molecule with a specific marker [5, 6]. Moreover, the devices used in optical detection are large and heavy, limiting the practical use of them. As an alternative method, electrochemical detection is used in biosensors. In general, electrochemical methods are thought to be more advantageous than optical methods due to reasons such as rapid response, cheap production, simple and easy to use, micro-suitability for biosensor production and low cost equipment for signal conversion [7-9]. Using a simple electrode, the oxygen consumption of the cells, the activity of some released proteins, or the analysis of cell proteins can easily be done [10, 11]. The ability of developing the technology to produce ultrafine and nano-level electrodes makes it possible to use these electrodes in micro and nano-level analyzes [12-14]. Unlike millimeter-sized electrodes, micro- or nanoelectrodes have features that enhance sensitivity, such as low ohmic potential drop, low double layer charging current and high molecular transport [15, 16]. In addition, small size electrodes enable analysis in small volume media.

Carbon fibers are highly conductive and due to their small dimensions, it is possible to produce micro and nano probes with them, which offers certain advantages such as the use of these electrodes in single cell analysis. Carbon fiber based microelectrodes are commonly used for local and highly sensitive detection of neurotransmitters such as dopamine or recording of neuronal action potentials known as spikes, enabling electrochemical monitoring of neurochemical activity of brain [17-20]. Carbon fiber electrodes (CFEs) can allow for electrochemical monitoring in short-time domains and recording of neuronal activity in real time [21, 22]. Tissue damage during monitoring with CFEs is quite limited compared to other tools such as microdialysis probes. Brain has a very complex structure, and tools like CFEs have helped scientists gain a substantial amount of knowledge regarding function in a very easy manner. In addition to being used in such measurements, CFEs also have potential for detection of biologically important molecules with great sensitivity and selectivity by simply modifying the surface of the electrode with a bio-recognition molecule [23-25]. In a recent study, Salazar and his colleagues modified CFEs with Prussian Blue to realize a glutamate microbiosensor for neuroscience applications [26]. They first modified the surface of CFEs with Prussian Blue using electrodeposition. Next, they coated the surface with poly-o-phenylenediamine (PoPD) for higher stability and polyethyleneimine (PEI) as a glucose oxidase (GOx)

Received on: 19.01.2017
Accepted on: 13.03.2017

immobilization agent. Using this microbiosensor, they successfully detected glutamate at a significantly low concentration (<50 µM). In another study, Lee and coworkers used CFEs for detection of glucose [27]. They first treated carbon fiber electrodes with KOH for activation, a process that yielded improved adsorption of GOx, then used urea treatment to improve the conductivity of the electrodes. According to the results, the sensitivity value of the modified electrodes was two to three times higher than untreated electrodes. These are some of the many studies showing the benefits of using CFEs for bio-sensing applications. In this study, the surface of CFEs were modified with Pt nanoparticles to detect O_2 and H_2O_2. Detection of cells' O_2 consumption is highly desirable as it plays a vital role in respiration and cell metabolism. It is also crucial for a variety of practical applications including steel-making and food preparation [28]. Detection of H_2O_2 is also important in a variety of fields ranging from food preparation to environmental monitoring [29, 30]. Moreover, the excess presence of H_2O_2 in cells causes oxidative stress and hereby leading to various diseases including cancer. Usually enzymes such as horse radish peroxidase are used to make H_2O_2 biosensors, however, complex fabrication procedures, low reproducibility and enzyme instability limits large scale application of these biosensors [5, 31]. Recently, more focus has been given to constructing enzymeless H_2O_2 biosensors and their practical applications. Basically, in this study, production of CFE probes was carried out by pulling carbon fiber inserted glass capillary tubes using a micropuller to seal single carbon fibers with a thin layer of glass. Subsequently, the electrode surface was ground to expose the disk electrode which then was modified with platinum nanoparticles. Lastly, the ability of these probes in oxygen and hydrogen peroxide detection was analyzed.

2. Materials and Methods

2.1. Materials

Patch clamp glass (OD/ID: 1.65/1.1 mm) (World Precision Instruments, USA), silver paste (Sigma-Aldrich, USA), chloroplatinic acid – H_2PtCl_6 (Sigma-Aldrich, USA), hydrogen peroxide solution, 30 % (w/w) in H_2O (Sigma-Aldrich, USA), phosphate buffered saline (PBS) (Sigma-Aldrich, USA), hydrogen chloride (HCl) (Sigma-Aldrich, USA), ferrocenemethanol (FMA) 97% (Sigma-Aldrich, USA).

Carbon fibers with a diameter of ~ 6 µm were kindly provided by Dr. Mustafa Erol (Izmir Katip Celebi University, Turkey)

2.2. CFE Fabrication and Characterization

First, carbon fibers were attached to copper wires that are only exposed at the ends using a silver (Ag) paste and then the connection was made permanent by baking and solidifying the silver paste at 180 °C. The carbon fibers attached to the copper wires were placed in the glass capillary tubes with the help of the copper wires, and the copper wires were fixed on the glass tubes with heat-shrinking rollers to prevent any damage to the carbon fibers in the subsequent operation. In the next process, a thin layer of glass was formed on carbon fibers by pulling the glass capillary tubes using micropulling method to produce CFEs (Figure 1A). PC-10 micropulling machine (Narishige, Japan) was used for micropulling. The pulling parameters were optimized prior to CFE fabrication. The parameters used for micropulling are as follows; option: 1 and heating level: 65. As the last step of the probe production, the tip of the CFEs were ground to expose the disk electrodes using a machine called microgrinder (EG-401, Narishige, Japan) (Figure 1B, C_I).

Following the production of CFEs, the electrochemical behavior of these probes was analyzed with a PBS solution containing electroactive ferrocenemethanol (FMA). PBS solution containing 1 mM FMA was prepared for the electrochemical analysis. Cyclic voltammetry (CV) of the CFEs were obtained in this solution using a potentiostat (Autolab PGSTAT204, Metrohm, Switzerland). At this point, the current was measured by sweeping the potential of the working electrode between 0 and +0.5 V (vs Ag / AgCl) at a scan rate of 50 mV / s.

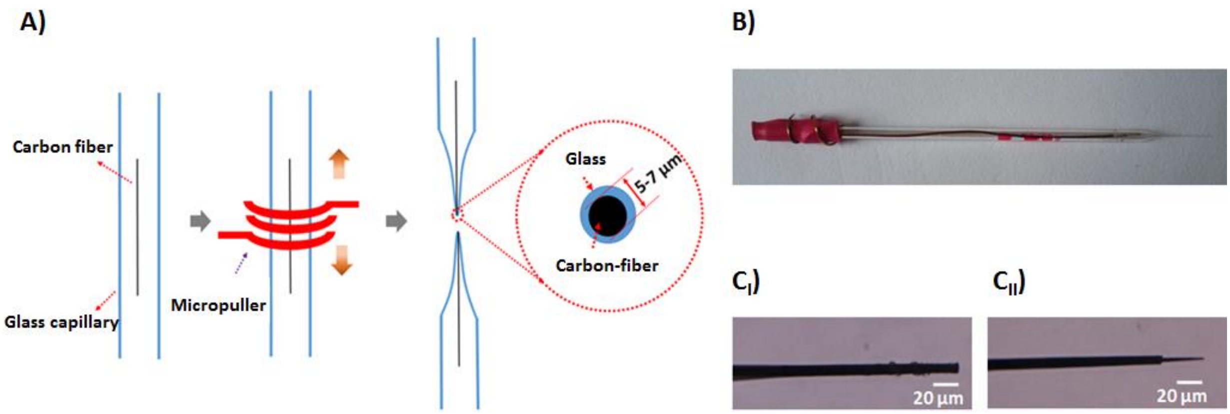


Figure 1. Fabrication of carbon fiber electrode (CFE) probes by micropulling carbon fiber inserted glass capillaries (A). Optical images of CFEs with disk (C_I) and cylinder (C_{II}) tips.

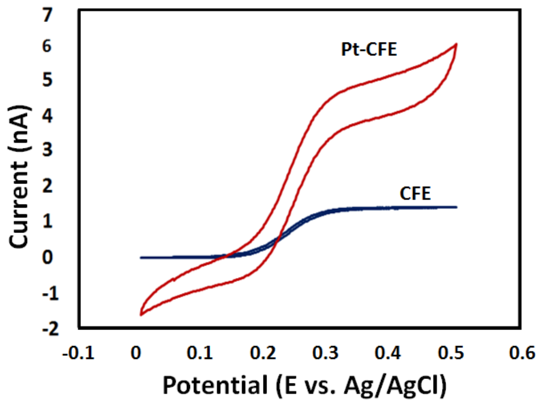


Figure 2. CV curves of a platinum-modified (Pt-CFE) and a bare CFE in 1 mM FMA + PBS.

2.3. Electrode surface Modification and Characterization

Bare CFEs were cleaned in acetone and milli-Q water prior to surface modification with Pt nanoparticles. Chronoamperometry was used for electrochemical modification. Basically, CFEs were placed in a solution containing 2 mM H_2PtCl_6 + 0.1M HCl and the potential was held at 0 V (vs. Ag / AgCl) for about 70 s.

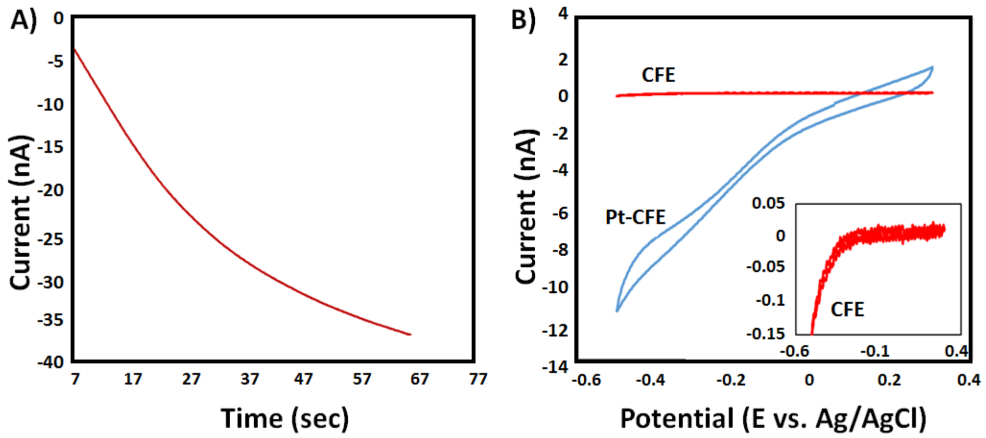


Figure 3. Modification of Pt-CFE with platinum nanoparticles in 0.2 mM H_2PtCl_6 + 0.1 M HCl at 0 V (vs. Ag/AgCl) (A) and comparison of the oxygen reduction potential of these electrodes with bare CFEs in PBS solution using CV (B).

Afterward, electrochemical behavior of Pt nanoparticle modified CFEs (Pt-CFEs) were analyzed in 1 mM FMA + PBS solution using CV, for which the potential of the working electrode was swept between 0 and +0.5 V (vs Ag / AgCl) at 50 mV / s.

2.4. Electrochemical Detection

CV was used for the electrochemical detection (reduction) of oxygen. Basically, bare and Pt-CFEs were placed in a PBS solution, respectively and the CV curves were obtained, at which time current was measured by sweeping the potential of the working electrode between +0.4 V and -0.6 V (vs. Ag / AgCl).

Next, hydrogen peroxide (H_2O_2) was electrochemically detected with Pt-CFEs and the detection limit of these probes was determined. At this point, CV curves for H_2O_2 at different concentration levels were obtained. The potential of the modified electrodes were swept between 0 V and +1.4 V (vs. Ag / AgCl) at a scan rate of 100 mV / s to obtain the respective CV curves. Chronoamperometry was used to determine the lowest concentration level that the Pt-CFE could measure. Basically, the Pt-CFE was placed in a PBS solution and during which the potential of the modified electrode was kept constant at +0.8 V (vs. Ag / AgCl). Once the current gained a steady state, concentrated solutions of H_2O_2 were added to realize first 1μM and then 10μM of H_2O_2 in PBS.

3. Results and Discussion

The leak-proof CFEs whose electrochemical behavior can be quantitatively analyzed, can be easily produced by the production method proposed and developed in this study. Briefly, carbon fibers were first connected to coated copper wires that are only exposed at the ends, because of several reasons such as carbon fibers need to be inserted into glass capillary tubes and thereafter connected to the electrochemical system. Then, carbon fibers fixed to copper wires were inserted into glass capillaries, which were then pulled using a micropuller to seal carbon fibers with a thin layer of glass. In the final step, the tip of the CFEs were ground to realize microdisk electrodes at the tip. The optical microscopy image of a CFE was taken after grinding its tip (Figure 1B, C_I). As can be understood from the pictures taken, the probe tip was completely flattened and the disk electrode was successfully produced. It is also possible to produce cone-shaped CFEs of various sizes with another machine called "microforge" without using the micro-grinding machine (Figure $1C_{II}$). Basically, flame etching is used to give carbon fiber at the tip of the CFE a cone-shape. In this study, only disk shaped electrodes were used for electrochemical detection. Following the production of CFEs, the electrochemical behavior of these probes was analyzed with a PBS solution containing electroactive FMA. PBS solution containing 1 mM FMA was prepared for the electrochemical analysis. CV of the CFEs were obtained in this solution using a potentiostat (Metrohm-Autolab PGSTAT204) (Figure 2).

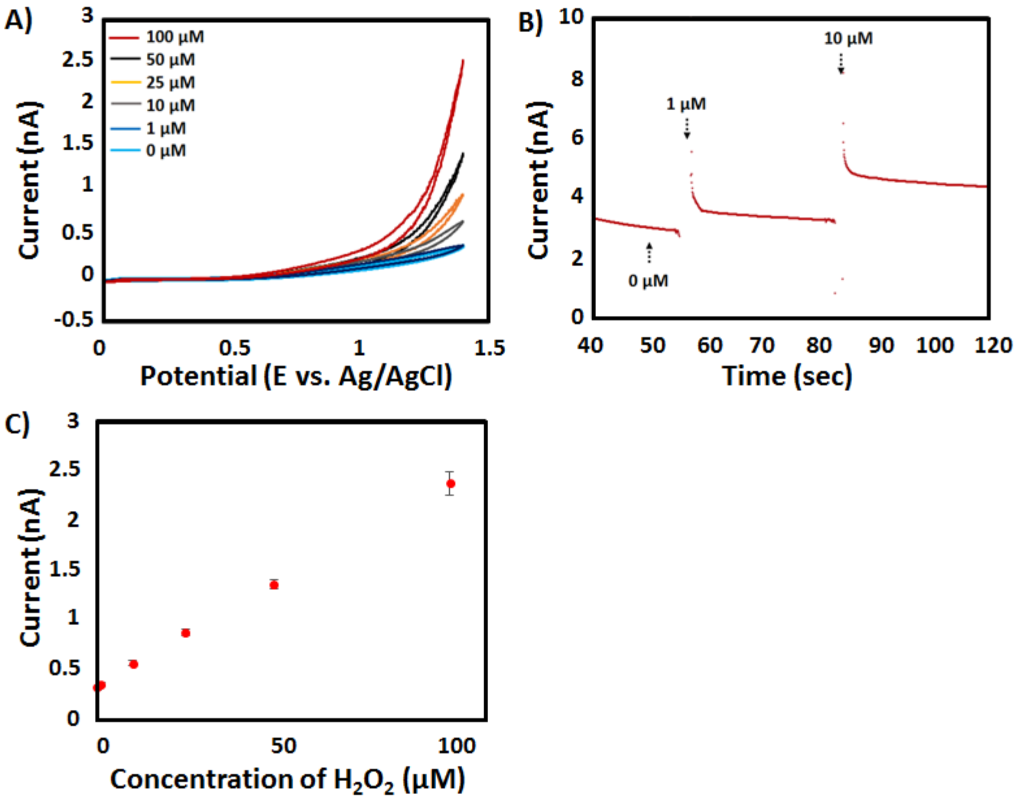


Figure 4. CV was used for electrochemical hydrogen peroxide detection with Pt-CFE (0, 1, 10, 25, 50 and 100 µM) (A). The smallest concentration level (limit of detection - LOD) that these electrodes can measure is determined by chronoamperometry (B). H_2O_2 calibration curve whose data were retrieved from respective CV curves (C).

At this point, the current was measured by sweeping the potential of the working electrode between 0 and +0.5 V (vs Ag / AgCl) at a scan rate of 50 mV / s. The maximum current observed in CV was around 1.5 nA, which is close to the theoretically calculated value. In the theoretical maximum current calculation, the formula "$I = 4nFDCr$" was used for the microdisk electrode (I: current, n: number of electrons (FMA = 1); F: Faraday constant (96485.329 s / mole); D: diffusion constant (D_{FMA}= 6.7x10-6 cm^2s^{-1}) C: analyte concentration (FMA = 1 mM); r: electrode radius (1 µm). This finding has led to the conclusion that the produced CFEs can be used for quantitative electrochemical detection. The surfaces of the CFEs in the next stage were electrochemically modified with platinum nanoparticles. Chronoamperometry was used for electrochemical modification (Figure 3A). During the process, the potential of Pt-CFE was held at 0 V (vs. Ag / AgCl). As can be understood from Figure 3A, the reduction current increased in negative direction during the electrochemical modification process as expected. It was then determined whether the modification process was successful by comparing the performance of a modified Pt-CFE with that of a normal CFE in both oxidizing FMA and reducing oxygen. First, a Pt-CFE was immersed in a solution containing 1 mM electroactive FMA + PBS in which case the potential of the electrode was swept from 0 V to 0.5 V vs. Ag/AgCl at a scan rate of 50 mV / s to obtain a CV curve. When compared with that of bare CFE, the peak current of Pt-CFE was significantly high in FMA solution (Figure 2). Higher current response indicates that the modification of CFEs with Pt nanoparticles increases the electrode surface area. The increase in the surface area of the electrode contributes positively to the sensitivity of such electrodes. Second, the performance of Pt-CFE in reducing of O_2 was determined and compared with that of bare CFE. Platinum has a high catalytic activity in the reduction of oxygen. CV was used for the electrochemical reduction of oxygen, during which the current was measured by sweeping the potential of the working electrode (carbon fiber and platinum modified carbon fiber electrodes) between +0.4 V and -0.6 V (vs. Ag / AgCl) at a scan rate of 50 mV / s. The reduction current obtained from platinum-modified electrodes was observed to be much higher than that of bare CFEs, indicating the successful modification of electrode surface with platinum nanoparticles (Figure 3B).

Next, the performance of Pt-CFEs in electrochemical detection of hydrogen peroxide (H_2O_2) was determined. At this point, the detection limit was also investigated. As can be understood from Figure 4A, the indicated H_2O_2 concentration levels (10, 25, 50, 100 µM) were successfully measured with the modified electrodes. Chronoamperometry was used to determine the lowest concentration level that the Pt-CFE could measure, and during this process the potential of the modified electrode was kept constant at +0.8 V (vs. Ag / AgCl). As can be concluded from these results, H_2O_2 can be measured at 1 µM with Pt-CFEs (Figure 4B). Considering the possibility of high potential damage to the electrode, +0.8 V (vs. Ag / AgCl) was used instead of +1.2 V (vs. Ag / AgCl). In the last step, a calibration curve was generated by using 3 different data for the respective concentration levels. According to the results, the Pt-CFEs showed a linear response between 0 and 100 μM H_2O_2 concentration levels (Figure 4C).

5. Conclusions

In this study, low-cost and highly sensitive CFE based O_2 and H_2O_2 sensing probes were fabricated. To the best of my knowledge, this is the first study where probe type ultra-micro CFEs were modified with Pt nanoparticles for such an application. Briefly, CFE probes were produced by micropulling carbon fiber inserted glass capillaries. After grinding the tip of the probe to realize microdisk electrode at the tip, the surface was electrochemically modified with Pt nanoparticles. Results indicate that O_2 and H_2O_2 were successfully detected using the probes. The introduced fabrication method is quite simple and cost efficient as carbon fibers are a lot cheaper than noble metals such as Pt or Au. In addition, the produced Pt-CFE probes have high potential for local analysis because of their small size and development of various microbiosensors with small surface modifications (e.g., glucose biosensor). Future work is planned for the use of Pt-CFEs in the development of a very sensitive ATP biosensor.

5. Acknowledgement

This study was supported by Scientific and Technological Research Council of Turkey (TUBITAK), (project no: 115C013).

6. References

[1] A. Erdem, G. Congur, "Label-free voltammetric detection of MicroRNAs at multi-channel screen printed array of electrodes comparison to graphite sensors", *Talanta, 118*, 7-13, **2014**.

[2] M. Sen, K. Ino, H. Shiku, T. Matsue, "Accumulation and detection of secreted proteins from single cells for reporter gene assays using a local redox cycling-based electrochemical (LRC-EC) chip device", *Lab on a Chip, 12* (21), 4328-4335, **2012**.

[3] M. Kaplan, T. Kilic, G. Guler, J. Mandli, A. Amine, M. Ozsoz, "A novel method for sensitive microRNA detection: Electropolymerization based doping", *Biosens Bioelectron*, **2016**.

[4] J. Narang, C. Singhal, N. Malhotra, S. Narang, A. K. Pn, R. Gupta, R. Kansal, C. S. Pundir, "Impedimetric genosensor for ultratrace detection of hepatitis B virus DNA in patient samples assisted by zeolites and MWCNT nano-composites", *Biosens Bioelectron, 86*, 566-74, **2016**.

[5] K. Y. Inoue, M. Matsudaira, R. Kubo, M. Nakano, S. Yoshida, S. Matsuzaki, A. Suda, R. Kunikata, T. Kimura, R. Tsurumi, T. Shioya, K. Ino, H. Shiku, S. Satoh, M. Esashi, T. Matsue, "LSI-based amperometric sensor for bio-imaging and multi-point biosensing", *Lab on a Chip, 12* (18), 3481-3490, **2012**.

[6] K. Ino, Y. Kanno, T. Nishijo, T. Goto, T. Arai, Y. Takahashi, H. Shiku, T. Matsue, "Electrochemical detection for dynamic analyses of a redox component in droplets using a local redox cycling-based electrochemical (LRC-EC) chip device", *Chem Commun (Camb), 48* (68), 8505-7, **2012**.

[7] T. G. Drummond, M. G. Hill, J. K. Barton, "Electrochemical DNA sensors", *Nat Biotechnol, 21* (10), 1192-1199, **2003**.

[8] J. Wang, "From DNA biosensors to gene chips", *Nucleic Acids Res*, *28* (16), 3011-3016, **2000**.

[9] G. S. Bang, S. Cho, B. G. Kim, "A novel electrochemical detection method for aptamer biosensors", *Biosens Bioelectron*, *21* (6), 863-870, **2005**.

[10] H. Shiku, T. Shiraishi, H. Ohya, T. Matsue, H. Abe, H. Hoshi, M. Kobayashi, "Oxygen consumption of single bovine embryos probed by scanning electrochemical microscopy", *Anal Chem, 73* (15), 3751-8, **2001**.

[11] M. Sen, K. Ino, H. Shiku, T. Matsue, "A new electrochemical assay method for gene expression using HeLa cells with a secreted alkaline phosphatase (SEAP) reporter system", *Biotechnol Bioeng, 109* (8), 2163-7, **2012**.

[12] Y. Takahashi, A. I. Shevchuk, P. Novak, Y. Zhang, N. Ebejer, J. V. Macpherson, P. R. Unwin, A. J. Pollard, D. Roy, C. A. Clifford, H. Shiku, T. Matsue, D. Klenerman, Y. E. Korchev, "Multifunctional nanoprobes for nanoscale chemical imaging and localized chemical delivery at surfaces and interfaces", *Angew Chem Int Ed Engl, 50* (41), 9638-42, **2011**.

[13] M. Sen, Y. Takahashi, Y. Matsumae, Y. Horiguchi, A. Kumatani, K. Ino, H. Shiku, T. Matsue, "Improving the electrochemical imaging sensitivity of scanning electrochemical microscopy-scanning ion conductance microscopy by using electrochemical Pt deposition", *Anal Chem, 87* (6), 3484-9, **2015**.

[14] P. Actis, S. Tokar, J. Clausmeyer, B. Babakinejad, S. Mikhaleva, R. Cornut, Y. Takahashi, A. Lopez Cordoba, P. Novak, A. I. Shevchuck, J. A. Dougan, S. G. Kazarian, P. V. Gorelkin, A. S. Erofeev, I. V. Yaminsky, P. R. Unwin, W. Schuhmann, D. Klenerman, D. A. Rusakov, E. V. Sviderskaya, Y. E. Korchev, "Electrochemical nanoprobes for single-cell analysis", *ACS Nano, 8* (1), 875-84, **2014**.

[15] R. Lin, P. L. Taberna, J. Chmiola, D. Guay, Y. Gogotsi, P. Simon, "Microelectrode Study of Pore Size, Ion Size, and Solvent Effects on the Charge/Discharge Behavior of Microporous Carbons for Electrical Double-Layer Capacitors", *J Electrochem Soc, 156* (1), A7-A12, **2009**.

[16] J. X. Wang, T. E. Springer, R. R. Adzic, "Dual-pathway kinetic equation for the hydrogen oxidation reaction on Pt electrodes", *J Electrochem Soc*, *153* (9), A1732-A1740, **2006**.

[17] I. M. Taylor, E. M. Robbins, K. A. Catt, P. A. Cody, C. L. Happe, X. T. Cui, "Enhanced dopamine detection sensitivity by PEDOT/graphene oxide coating on in vivo carbon fiber electrodes", *Biosens Bioelectron*, **2016**.

[18] F. Vitale, S. R. Summerson, B. Aazhang, C. Kemere, M. Pasquali, "Neural stimulation and recording with

bidirectional, soft carbon nanotube fiber microelectrodes", *Acs Nano, 9* (4), 4465-74, **2015**.
[19] G. Guitchounts, J. E. Markowitz, W. A. Liberti, T. J. Gardner, "A carbon-fiber electrode array for long-term neural recording", *J Neural Eng, 10* (4), 046016, **2013**.
[20] L. Qi, E. L. Thomas, S. H. White, S. K. Smith, C. A. Lee, L. R. Wilson, L. A. Sombers, "Unmasking the Effects of L-DOPA on Rapid Dopamine Signaling with an Improved Approach for Nafion Coating Carbon-Fiber Microelectrodes", *Anal Chem, 88* (16), 8129-36, **2016**.
[21] M. L. Huffman, B. J. Venton, "Carbon-fiber microelectrodes for in vivo applications", *Analyst, 134* (1), 18-24, **2009**.
[22] M. D. Nguyen, B. J. Venton, "Fast-scan Cyclic Voltammetry for the Characterization of Rapid Adenosine Release", *Comput Struct Biotechnol J, 13*, 47-54, **2015**.
[23] W. H. Oldenziel, B. H. Westerink, "Improving glutamate microsensors by optimizing the composition of the redox hydrogel", *Anal Chem, 77* (17), 5520-8, **2005**.
[24] W. H. Oldenziel, G. Dijkstra, T. I. Cremers, B. H. Westerink, "Evaluation of hydrogel-coated glutamate microsensors", *Anal Chem, 78* (10), 3366-78, **2006**.
[25] X. Lin, X. Jiang, L. Lu, "DNA deposition on carbon electrodes under controlled dc potentials", *Biosens Bioelectron, 20* (9), 1709-17, **2005**.
[26] P. Salazar, M. Martín, R. D. O'Neill, J. L. González-Mora, "Glutamate microbiosensors based on Prussian-Blue modified carbon fiber electrodes for neuroscience applications: In-vitro characterization", *Sensors and Actuators B: Chemical, 235*, 117-125, **2016**.
[27] J.-H. Kim, S. Cho, T.-S. Bae, Y.-S. Lee, "Enzyme biosensor based on an N-doped activated carbon fiber electrode prepared by a thermal solid-state reaction", *Sensors and Actuators B: Chemical, 197*, 20-27, **2014**.
[28] N. Dossi, R. Toniolo, A. Pizzariello, E. Carrilho, E. Piccin, S. Battiston, G. Bontempelli, "An electrochemical gas sensor based on paper supported room temperature ionic liquids", *Lab Chip, 12* (1), 153-8, **2012**.
[29] J. Tian, Q. Liu, C. Ge, Z. Xing, A. M. Asiri, A. O. Al-Youbi, X. Sun, "Ultrathin graphitic carbon nitride nanosheets: a low-cost, green, and highly efficient electrocatalyst toward the reduction of hydrogen peroxide and its glucose biosensing application", *Nanoscale, 5* (19), 8921-4, **2013**.
[30] W. Chen, S. Cai, Q. Q. Ren, W. Wen, Y. D. Zhao, "Recent advances in electrochemical sensing for hydrogen peroxide: a review", *Analyst, 137* (1), 49-58, **2012**.
[31] D. Liu, T. Chen, W. Zhu, L. Cui, A. M. Asiri, Q. Lu, X. Sun, "Cobalt phosphide nanowires: an efficient electrocatalyst for enzymeless hydrogen peroxide detection", *Nanotechnology, 27* (33), 33LT01, **2016**.

Dr. Mustafa ŞEN received his M.Sc. degree in molecular biology and genetics from Istanbul Technical University in 2010 and PhD degree in bio-enginnering from Tohoku University, Japan in 2013. After spending one and a half year as a postdoctoral researcher at Tohoku University, he accepted a position at Izmir Katip Celebi University, where he was appointed as an Assist. Professor. His main research focus is in the field of micro- and nanobiosensor development and their applications in single cell and micro-tissue analysis.

EVALUATION OF STRESS PARAMETERS BASED ON HEART RATE VARIABILITY MEASUREMENTS

Fatma UYSAL[1], Mahmut TOKMAKÇI[1]

[1]Department of Biomedical Engineering, Erciyes University, Kayseri, Turkey
fatmauysal@erciyes.edu.tr, tokmakci@erciyes.edu.tr

Abstract: *In this study, heart rate variability measurements and analyses were carried out with help of the ECG recordings to show how autonomic nervous system activity changes. ECG signal recording was collected from six volunteer participants under following conditions; the situation of relaxation, Stroop color/word test, mental test and auditory stimulus in a laboratory environment so as to evaluate the parameters related to stress. Totally seven minutes ECG recording was taken and analyzed in time and frequency domain. We investigated that how autonomic nervous system activity changed in the position of stress. According to frequency domain analysis, there has seen significant increase at low frequency band (LF band) which reflects sympathetic activity particularly during Stroop color/word and mental tests.*
Keywords: *Stress, heart rate variability, autonomic nervous system*

1. Introduction

Stress which is a part of daily routine is body's reaction to sensed emotional, mental or physical affliction [1]. Some studies [2, 3] showed that reasonable levels of stress can provide positive effect such as an increment in concentration power. Extreme levels of stress has destructive effects such as cognitive dysfunctions, depression, cardiovascular diseases [4,5] or psychiatric disorders such as anxiety, depression, and Alzheimer [6]. The autonomic nervous system (ANS) which performs dynamically regulation of our body system through sympathetic (crucial in danger situations) and parasympathetic (induces relaxation response) branches is triggered under stress condition. In this situation, parasympathetic nervous system is depressed while sympathetic nervous system is activated. The activation of sympathetic nervous system leads to increase of blood pressure and heart rate [7]. The changes that occur at heart rate are called heart rate variability (HRV). In other words, HRV is time differences between two consecutive R waves [8]. HRV is assessed either by time domain or frequency domain analyses which are including power spectral density analysis. HRV is disposed to be higher and more complex when the ANS is in balance. Otherwise, HRV tends to be lower. This situation can be used for stress evaluation.

Electrocardiogram signal (ECG) is accepted gold standard in HRV analysis [9]. Moreover, ECG signal recording reflects the electrical activity of the heart which is controlled by the ANS.

Numerous studies [9-13] have been performed to understanding and evaluation of stress during the last few decades. HRV which is derived from photopletismograph or electrocardiogram (ECG) signal was analyzed in these studies. HRV analyses which are performed both time and frequency domain are calculated by time differences between consecutive R waves in the ECG signal.

There are various studies [7,8,11,14-16] using different combinations of stress conditions to calculate HRV based on ECG signals under stress and stress free situations. Taelman et al. [7] studied the effect of mental stressor on HRV and reported that a significant difference between the rest and mental task conditions. The relationship between HRV and stress generated by visual stimuli has been investigated. Significant changes in HRV that occurred in the case of visual stress was determined [14,15]. Oh et al. [16] investigated the effects of noisy sounds on human stress by using ECG signals. The results of these signals showed that some of these noisy sounds lead to increase the stress level on humans. Visnovcova et al. [11] observed a reduction of HRV complexity under two different stress situations (Stroop test, mental arithmetic) with regard to the baseline. A large part of studies demonstrated that standard deviation of all NN interval (SDNN) and high frequency band (HF) decreased whereas LF and LF/HF increased during stress situations [8].

The main purpose of this study is to examine whether there is a relationship between stress and HRV parameters which are derived from ECG signals and determine the

Received on: 28.01.2017
Accepted on: 13.03.2017

effect of stress on ANS activity. We also investigated how auditory stimulus has an effect on the ANS activity.

The remainder of this paper is organized as follows: data collection, experimental protocol and signal analysis are explained in Section 2. In Section 3 and Section 4, calculated HRV parameters and discussion are ensured respectively. Finally, conclusion is presented in Section 5.

2. Method

2.1. Participants

A total of 6 volunteer participants 2 of female and 4 of male with mean age of 27 (±3.87) took part in our experimental study. All participants were asked to drink tea, coffee, alcohol, etc. which affect the cardiovascular system for three hours before the experiment. None of the subjects had no history of cardiovascular disease and psychiatric condition.

2.2. Data Collection

ECG signals were acquired by using Biopac MP36 unit at a sampling rate of 1000 Hz sampling frequency. The subjects sat on a comfortable chair and were verbally informed about the aim and procedure of this study. For ECG signal recording, Ag/AgCl electrodes were placed at subject's right forearm (negative), left leg (positive) and right leg (ground) to make possible recording of Lead II trace. ECG signal recording was performed in the following situations: baseline (S1), Stroop color/word test (S2), post-stress recovery (S3), mental test (S4), recovery (S5), auditory stimulus (S6), recovery (S7). We aimed to induce stress at the subject using Stroop color/word test, mental test and auditory stimulus situations. To be able to make an equitable comparison, each task session had a same duration of one minute. The procedure for the experiment is shown in Fig.1.

Firstly, one minute signal was received from each participant and this signal was accepted baseline. ECG signal was collected without any stimulus for one minute in this part. After this section, various tests are applied to induce stress at the subject.

2.3 Stress Conditions

2.3.1. Stroop Color/Word Test

Stroop color/word test [17] is a cognitive test which was carried out by J.R. Stroop in 1953. The test consists of three parts, but only the third part has been applied to the participants. Participants are asked to read the words written in different colors from the color name printed on the paper as quickly as possible. The result is known as Stroop effect that subjects are forced to read color names written in different color.

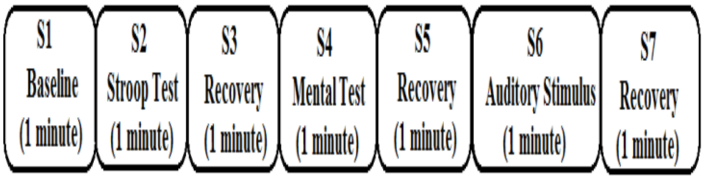


Fig. 1. Procedure of the experiment

Stroop color/word test was applied to the subjects for one minute. After that recovery session (S3) was performed for heart rate stabilization and rest after stress. Fig. 2 shows a part of the Stroop color/word test.

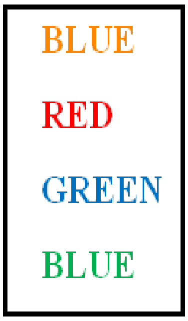


Fig. 2. A part of Stroop color/word test

2.3.2 Mental Test

Mental arithmetic test is one of the methods frequently used in the literature to induce stress [9, 11, 18, 19]. Participants were asked to perform arithmetic task for 1 minute. Briefly, subjects started to subtract 7 then subtract 7 and so on, starting from 3000 as quickly as possible. Results showed that all participants were stressed in this section.

2.3.3 Auditory Stimulation

Auditory stimulus can be used as stress inducer [16]. In this section participants were exposed to tension musics for 1 minute via an earphone with their eyes closed. It has been said that at any moment a needle immersed in order to increase the stress level of participants.

2.4 Signal Analysis

After ECG recordings were taken the signal was filtered to remove noise and baseline drift before HRV analysis. Firstly, moving average filter was used to eliminate noise for filtering after that a high pass FIR filter with a 0.5 Hz cut of frequency was used to remove baseline drift. Lastly notch filter was applied to eliminate 50 Hz power line interference.

Fig. 3 shows flowchart which is representing the process steps for HRV analysis. Baseline shift in the signal was removed after filtering and noise was eliminated as shown in Fig. 4.

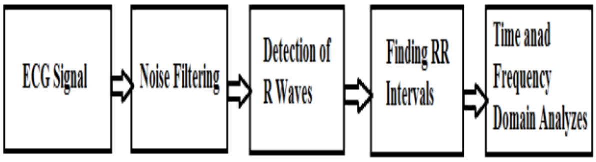


Fig. 3. Process steps for ECG Signal Analysis

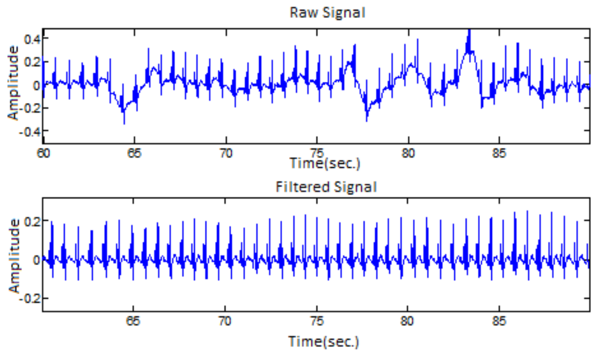


Fig. 4. Raw and filtered ECG signal

R wave detection in the ECG signal was performed by using the Pan Tompkins algorithm [20] as shown in Fig. 5. Successful inter-beat intervals (between two R waves) are calculated by using HRV analyses and these analyses are carried out in two steps which are time and frequency domain analyses.

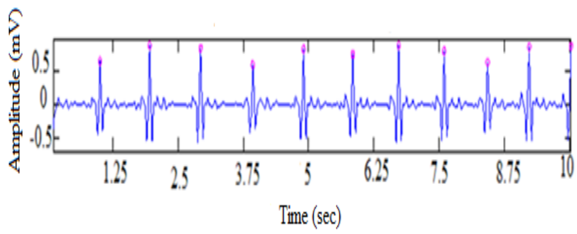


Fig. 5. R wave detected signal

Measurement standards which are recommended by the European Society of Cardiology and North American Society of Pacing and Electrophysiology guidelines (ESC / NASPE) are accepted for HRV analyses [21]. Some significant time and frequency domain parameters and descriptions are shown in Table 1. As seen in Table 1 time domain measurements include the mean of RR intervals (Mean RR), standard deviation of normal normal intervals (SDNN), the square root of mean squared difference of successive RR's (RMSSD) and percentage of normal normal intervals that vary by more than 50 milliseconds from the previous interval (PNN50). Frequency domain HRV measurements contain very low frequency power (VLF), low frequency power (LF), high frequency power (HF) and LF/HF. There is a need for time series which are calculated from the consecutive RR intervals, in the time domain measurements. Frequency domain measurements and power spectral density (PSD) are calculated by using fast Fourier transformation. Time domain methods are very simple to calculate because they are implemented directly to the series of consecutive RR interval values. However, these measurements do not provide information on the amount of autonomic balance or the temporal distribution of the power different branches of the ANS. The amount of autonomic balance can be determined at any given time with frequency domain methods.

Table 1. Time and frequency domain parameters for HRV analyses

Variable	Unit	Description
Time Domain Parameters		
Mean RR	ms	Average RR interval
SDNN	ms	Standard deviation of all NN interval
RMSSD	ms	The square root of the mean squared difference between adjacent N-N intervals, reflects mainly vagally influence
pNN50	%	Percentage of normal normal intervals greater than 50 milliseconds
Frequency Domain Parameters		
VLF	ms^2	Power spectrum band between 0.003-0.04 Hz
LF	ms^2	Power spectrum band between 0.04-0.15 Hz, reflects sympathetic activity
HF	ms^2	Power spectrum band between 0.15-0.4 Hz, reflects parasympathetic activity
LF/HF	n.u.	Reflects sympathovagal balance

3. Results

In this study, ECG signals during seven minutes received from volunteer participants under different test procedures were analyzed. The first minute of these signals was accepted as the reference (baseline). Results were interpreted by comparing the change of parameters in the stress condition with this reference signal.

The comparisons of HRV parameters under different conditions with mean and standard deviation values are shown in Table 2.(a) and 2.(b). The time and frequency domain parameters given in tables are the most commonly used parameters for measuring and evaluating stress in the literature. As shown in Table 2.(a) the mean value of RR intervals was 737.4±91.1 during the rest state, it considerably decreased 691±47,8 in Stroop color/word test and 686.2±64.9 in mental task. There was an important distinction between these situations and these values show that when the person feels under stress heart rate accelerates. Significant changes were observed in LF band which is reflecting sympathetic activity. LF power is 427.8±442.3 at the baseline level, while the power at this band in the Stroop and mental test is 813.4±1154 and 1432±1029 respectively. A significant increase in sympathetic activity has occured as expected in stress situations. Appropriately, the ratio of between LF and HF power which is reflecting sympathovagal balance increased due to the increase in LF power during Stroop and mental test. The auditory stimulus that leads to a person's tension state does not seem to cause a significant change in the level of stress of the person.

Result of the analysis on the frequency domain belongs to one of the participants is seen in Fig. 6. Figure 6.a and 6.b demonstrate power spectral density graphs of the person’s resting state and mental test respectively. While the power of the LF band is 580.1 ms^2 and the power of the HF band is 451.4 ms^2 in resting state, the power of the LF band increased to 1945.4 ms^2 and the power of the HF band decreased to 91.2 ms^2 during the mental test. Sympathetic activity and the power of the LF band reflecting the sympathetic activity increased during the mental test.

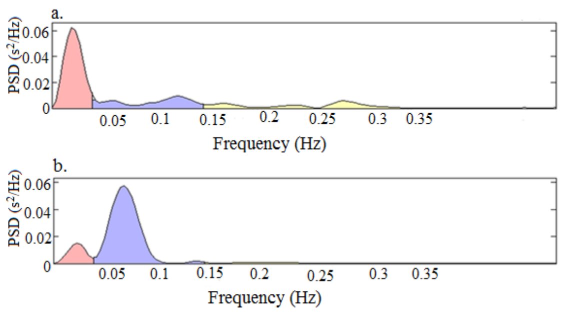


Fig. 6. Power spectral density graph of a participant during rest state (a) and mental test (b). (Variance is indicated s^2)

How Mean Normalized RR intervals of the participants have changed during resting state, Stroop color/word test, mental test and auditory stimulus is depicted in Fig. 7. A significant difference was not observed between the baseline and the auditory stimulus condition, whereas there was a decrease in average RR interval in almost all participants during the Stroop color / word and mental test.

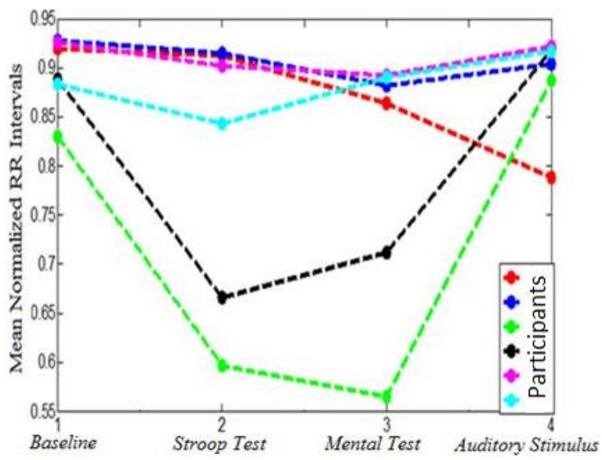


Fig. 7. Mean Normalized RR interval change in baseline, stroop test, mental activity and auditory stimulus states belong to 6 participants

Table 2.(a) Mean ± standard deviation values of time domain parameters under different conditions

	Mean RR	SDNN	RMSSD	PNN50	Mean HR
Baseline (S1)	737,4± 91,1	33,1± 11,1	20± 9,9	4,7± 6,4	82,5± 9,2
Stroop color/ word test (S2)	691± 47,8	37,8± 19	22± 13,5	6,6± 9,8	87,4± 5,6
Recovery (S3)	751,12± 78,7	45,3± 22,1	26,8± 14,3	7,6± 6,5	81± 7,8
Mental test (S4)	686,2± 64,9	45,7± 13,9	25,14± 12,2	8,4± 9,9	88,4± 7,4
Recovery (S5)	742,6± 66	53,67± 31	41,67± 27	9,05± 9,9	81,7± 6,7
Auditory stimulus (S6)	753,7± 102	29,7± 9,5	25,5± 11,4	6,2± 8,9	80,9± 10
Recovery (S7)	740± 68,8	42,1± 12,5	26,8± 10,5	6,3± 7	81,9± 7,1

Table 2.(b) Mean ± standard deviation values of frequency domain parameters under different conditions

	LF	HF	LF/HF
Baseline (S1)	427,8±442,3	271,8±275	1,8±1
Stroop color/word test (S2)	813,4±1154	287,4±275	3,7±3,4
Recovery (S3)	972±1377	312±374	3,5±4,7
Mental test (S4)	1432,8±1029	381±322	6,7±7,6
Recovery (S5)	1385±2200	538,4±768	2,7±2,27
Auditory stimulus (S6)	349±278,8	179±162,7	2,76±2,3
Recovery (S7)	836±675	295±255	7,9±11,1

4. Discussion

The present study explored the effect of different stressors (Stroop color/word test, mental test and auditory stimulus) on ANS activity by using HRV time and frequency domain analyses. We assessed the HRV parameters acquisition from ECG signal recording for two conditions: rest and different stress situations. According to the results obtained from study, a significant decrease was observed in the average RR interval in Stroop color/word and mental test. Mean heart rate (mean HR), SDNN and pNN50 values calculated by using time domain measurements increased during the Stroop color/word and mental test, these findings support the results in [19]. Unlike Stroop color/word and mental test there is no significant change in the auditory stimulus state. This finding demonstrates that the auditory stimulus does not cause a situation that would stress someone out. It was also observed that the meaningful changes in time and frequency domain occur in the recovery (S7), immediately after the auditory stimulus, rather than in the auditory stimulus (S6). A considerable increase was observed in the LF band reflecting the sympathetic activity during the Stroop color/word and mental test, when compared with resting state. However, unlike [9, 11] there is no significant difference in the HF power between rest and stress conditions except auditory stimulus situation. A slight decrease was detected in the HF band only in the case of auditory stimulus. These differences between rest and different stress conditions are expected and results are in agreement with the other studies [8]**.**

5. Conclusion

In this study we demonstrated clearly that HRV is a very useful tool to show how ANS activity changes under stress condition. We investigated HRV characteristics for rest and three different stress situations which are Stroop color/word test, mental test and auditory stimulus. According to time and frequency domain analyses some significant differences were observed between rest and stress conditions. From HRV data presented here, we conclude that under stress situation there was an increase in LF power which reflects sympathetic activity whereas a decrease in the HF band which reflects parasympathetic activity.

The system design will be performed by us to collect data besides this, the results will be made more general with the use of statistical analysis methods in future work of the study.

Acknowledgement

This work was supported by Research Fund of the Erciyes University. Project number: FYL-2016 6599.

6.References

[1] S. Reisman, "Measurement of physiological stress", In *Bioengineering Conference, 23rd Northeast,1997* pp. 21-23.

[2] K. Plarre, A. Raij, S. M. Hossain, A. A. Ali, M. Nakajima, M. Al'absi, D.Siewiorek, "Continuous inference of psychological stress from sensory measurements collected in the natural environment", In *Information Processing in Sensor Networks (IPSN),10th International Conference,* pp. *97-108*, 2011.

[3] B. S. Oken, I. Chamine, W. Wakeland, "A systems approach to stress, stressors and resilience in humans", *Behavioural Brain Research*, vol. *282*, pp. *144-154*, 2015.

[4] O. Kofman, N. Meiran, E. Greenberg, M. Balas, H. Cohen, "Enhanced performance on executive functions associated with examination stress: Evidence from task-switching and Stroop paradigms", *Cognition& Emotion*, vol. *20,* no *5*, pp. *577-595,* 2006.

[5] O. V. Crowley, P. S. McKinley, M. M. Burg, J. E. Schwartz, C. D. Ryff, M. Weinstein, R. P. Sloan, "The interactive effect of change in perceived stress and trait anxiety on vagal recovery from cognitive challenge", *International Journal of Psychophysiology*, vol. *82,* no *3*, pp. *225-232*, 2011.

[6] A. Hernando, J. Lázaro, E. Gil, A. Arza, J. M. Garzón, R. López-Antón, R. Bailón, "Inclusion of respiratory frequency information in heart rate variability analysis for stress assessment", *IEEE journal of biomedical and health informatics*, vol. *20,* no *4*, pp. *1016-1025,* 2016.

[7] J. Taelman, S. Vandeput, A. Spaepen, S. Van Huffel, "Influence of mental stress on heart rate and heart rate variability" In *4th European conference of the international federation for medical and biological engineering*, pp. *1366-1369,* Springer Berlin Heidelberg, 2009.

[8] R. Castaldo, P. Melillo, U. Bracale, M. Caserta, M. Triassi, L. Pecchia, "Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis", *Biomedical Signal Processing and Control*, vol *18*, pp. *370-377*, 2015.

[9] D. McDuff, S. Gontarek, R. Picard, "Remote measurement of cognitive stress via heart rate variability" *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. *2957-2960*, 2014.

[10] J. Taelman, S. Vandeput, E. Vlemincx, A. Spaepen, S. Van Huffel, "Instantaneous changes in heart rate regulation due to mental load in simulated office work", *European Journal of Applied Physiology*, vol *111,* no *7*, pp. *1497-1505*, 2011.

[11] Z. Visnovcova, M. Mestanik, M. Javorka, D. Mokra, M. Gala, A. Jurko, I. Tonhajzerova, "Complexity and time asymmetry of heart rate variability are altered in acute mental stress", *Physiological measurement*, vol. *35,* no *7*, pp. *1319*, 2014.

[12] S. Vandeput, J. Taelman, A. Spaepen, S. Van Huffel, "Heart rate variability as a tool to distinguish periods of physical and mental stress in a laboratory environment", In *Proceedings of the 6th international workshop on biosignal interpretation (BSI), New Haven, CT*, pp. *187-190*, June, 2009.

[13] B. Kaur, J. J. Durek, B. L. O'Kane, N. Tran, S. Moses, M. Luthra, V. N. Ikonomidou, "Heart rate variability (HRV): an indicator of stress", In *SPIE Sensing Technology+ Applications* pp. *91180V-91180V*, International Society for Optics and Photonics, May, 2014.

[14] W. Wu, J. Lee, "Development of full-featured ECG system for visual stress induced heart rate variability (HRV) assessment" In *The 10th IEEE International Symposium on Signal Processing and Information Technology*, pp. *144-149*, 2010.

[15] W. Wu, J. Lee, H. Chen, "Estimation of heart rate variability changes during different visual stimulations using non-invasive continuous ecg monitoring system", In *Bioinformatics, Systems Biology and Intelligent Computing, IJCBS'09. International Joint Conference on*, pp. *344-347*, 2009.

[16] B. S. Oh, Y. K. Yeo, F. Y. Wan, Y. Wen, Y. Yang, Z. Lin, "Effects of noisy sounds on human stress using ECG signals: An empirical study", In *2015 10th International Conference on Information, Communications and Signal Processing (ICICS)*, pp. *1-4*, 2015.

[17] P. Renaud,J. P. Blondin, "The stress of Stroop performance: Physiological and emotional responses to color–word interference, task pacing, and pacing speed", *International Journal of Psychophysiology*, vol *27* no *2*, pp. *87-97*, 1997.

[18] S. Mayya, V. Jilla, V. N. Tiwari, M. M. Nayak, R. Narayanan, "Continuous monitoring of stress on smartphone using heart rate variability", In *Bioinformatics and Bioengineering (BIBE),IEEE 15th International Conference on* pp. *1-5*, 2015.

[19] V. Vuksanović, V. Gal, "Heart rate variability in mental stress aloud", *Medical engineering & physics*, vol. *29,* no *3*, pp. *344-349*, 2007.

[20] J. Pan, W. J. Tompkins, "A real-time QRS detection algorithm", *IEEE transactions on biomedical engineering*, vol *3*, pp. *230-236*, 1985.

[21] Task Force of the European Society of Cardiology, "Heart rate variability standards of measurement, physiological interpretation, and clinical use", *Eur heart J*, vol. *17*, pp. *354-381*, 1996.

Note:

Mahmut Tokmakci received M.Sc. and Ph.D degree in Electronic Engineering from Erciyes University, 1996 and 2003 respectively. He is currently Associate Professor Department of Biomedical Engineering, Erciyes University. His research areas include biomedical instrumentation, clinical engineering and biosignal processing.

Fatma Uysal received B.Sc. and M.Sc. degree in Biomedical Engineering from Erciyes University, Kayseri, Turkey in 2014 and 2017 respectively. She is currently research assistant in Erciyes University, Department of Biomedical Engineering, Kayseri, Turkey. Her research interests include biosignal processing and biomedical instrumentation.

Determining of Gastroparesis Disease from Electrogastrogram Signals Using Cramer-Rao Lower Bound and Power Spectral Density

Çiğdem Gülüzar Altıntop[1], Fatma Latifoğlu[1], Emre Çelikzencir[1], Gülten Can Sezgin[2], Alper Yurci[2]

[1]Erciyes University, Dept. of Biomedical Engineering, Faculty of Engineering, 38039, Kayseri, Turkey
[2]Erciyes University, Dept. of Gastroenterology, Faculty of Medicine, 38039, Kayseri, Turkey
cigdemacer@erciyes.edu.tr, flatifoglu@erciyes.edu.tr, emrecelikzencir1@gmail.com, gcsezgin@yahoo.com.tr; yurci@erciyes.edu.tr

Abstract: *Gastroparesis is a chronic disease of stomach mobility, defined as delay in the emptying of food from the stomach without mechanical obstruction. Various methods are used for the diagnosis of gastroparesis and scintigraphy method is accepted as the gold standard. In this study, Electrogastrogram (EGG) signals were obtained from gastroparesis patients and healthy volunteers using cutaneous electrodes. Unlike the methods used in the frequency analysis of EGG signals in the literature, parametric methods have been used in this study and the selection of the method to be used has not been determined intuitively, it has been determined by mathematical calculations. Cramer-Rao Lower Bound (CRLB) method has been used to determine which method should be used for obtaining Power Spectral Density (PSD). Using selected parameter estimation method, PSD* functions *were obtained. Several features were extracted from the PSD* functions *and they were utilized to differentiate patient and healthy groups. As a result, features that can classify EGG signals from gastroparesis patients and healthy subjects, have been obtained. Best electrode placement that can be used for this disease has been achieved succesfully.*

Keywords: *Gastroparesis, Electrogastrogram, Cramer-Rao Lower Bound, Autoregressive Parameter, Power Spectral Density*

1. Introduction

Gastroparesis is a chronic disease of stomach mobility, defined as delay in the emptying of food from the stomach without mechanical obstruction. Scintigraphic imaging, endoscopy, gastric manometer, breath testing and wireless capsule methods are used for diagnosis. The widely used diagnostic method in the clinic is endoscopy and scintigraphy. Delayed gastric emptying, which confirms the suspicion of gastroparesis, is measured by gastric scintigraphy. Scintigraphic imaging is accepted as gold standard in diagnosing delayed gastric emptying. However, scintigraphy is an invasive method. The main disadvantage of this method is that radionuclide is given to the subjects to be diagnosed and they radiate radioactive substance for imaging by gamma camera. Also, scintigraphy is an expensive imaging method. These disadvantages lead to search for different diagnostic methods for diagnosis of gastroparesis. EGG is an non-invasive and easy-to-implement method but it isn't routinely implemented because its results couldn't be standardized. However, the interest in this area has increased recently [1-3].

The EGG is a measurement method that gives information about the electrical activity of stomach. The EGG is obtained using surface electrodes placed on the skin surface in the abdominal region. First studies about EGG were performed by Alvarez in 1922 [4]. EGG has become a more commonly used method in practice due to the fact that it is not harmful to the patient. However its clinical use is limited. EGG is used commonly in diagnosing stomach disease, in treatment, and in electro physical studies. Gastric Electrical Activity (GEA) consists of two activities: Electrical Control Activity (ECA) and Electrical Response Activity (ERA). The ECA potential is a slow wave, and the ERA potential is called spike potential. The ECA is an activity that always identifies the stomach contractions. The slow rate of change of this activity has caused the ECA activity to be called slow waves. The ERA, which occurs for a longer period than the ECA, provides information about contractions. The normal frequency of the stomach ECA has 0.03 Hz-0.07 Hz. frequency range and 2.4-3.7 cpm (cycle per minute). Arrhythmias that may occur in the stomach will cause changes in the frequency spectrum of the EGG signals. Frequency range of the EGG signal 0,0033 Hz-0,03 Hz. and 0.2-2.4 cpm range shows bradygastria, 0.07 Hz.-0.15 Hz. and 3,7-9,0 cpm range

Received on: 30.01.2017
Accepted on: 13.03.2017

shows tachygastria. Obtaining these dysrhythmias in the EGG analysis is believed that provide diagnostic benefits. Analysis of arrhythmias of the EGG is defined as a diagnostic method in patients suffering from impaired gastric motility, vomiting, nausea, bloating, early saturation and reflux [5].

Gastric cancer, reflux, gastritis, indigestion, stomach ulcers, epigastria burning , unexplained vomiting, gastroparesis (delayed gastric emptying) and stomach irritability are common in gastric diseases [6].

Electrogastrography isn't a widely used method for the diagnosis of gastroparesis, but provides detailed information on the pathogenesis of gastrointestinal disorders. JD Chen et al. determined that saturation EGG abnormalities may detect delayed gastric emptying and showed that gastric emptying was associated with the EGG [1-3].

In literature, the EGG signals are analyzed using methods such as Power Spectral Density (PSD), Wavelet Transform (WT), Discrete Wavelet Transform (DWT), Fourier Transform (FT), Short-time Fourier Transform (STFT) and Fast Fourier Transform (FFT). Studies generally aimed to obtain dysrhythmia. Often time-frequency or frequency analysis methods are used [7-9]. In these studies methods to obtain PSD functions were selected instinctively.

In this study, CRLB method has been used to determine which method should be used for obtaining PSD of EGG signals [10, 11]. The best parameter estimation method was chosen per CRLB method. PSD functions were obtained for patients and healthy subjects using EGG signals. For classification, features have been extracted from PSD functions. The classification success was measured by Receiver Operating Characteristic (ROC) curves.

2. Methodology

For electrode placement to obtain EGG signals, different points and leads were used in scientific studies because of there is no exact points considered as gold standard. This shows that there should be many studies that need to be done on the acquisition and analysis of EGG signals [10, 12].

2.1. Data Acquisition

Patients who were diagnosed with gastroparesis at Erciyes University Medical Faculty Hospital were included this study.

In this study, the EGG signals were obtained from volunteers for 30 minutes on an empty stomach and 30 minutes on a full stomach for a total of 1 hour duration. After the fasting record, 536 kcal cheese (or turkey) sandwiches and 200 ml fruit juice menu standardized by dietitian were given to volunteers and the postprandial measurements were obtained (menu including: 69.4% carbohydrate, 13.8% protein and 16.8% fat). The medications used by the patient were discontinued at least 3 days prior to the EGG procedure. The volunteers were asked to stop eating and drinking 8 hours before recording. The EGG recordings were obtained with 2 separate channels from the Biopac MP-150 system using 6 electrodes. First channel was obtained from 1st and 4th region's electrodes; second channel was obtained from 2nd and 3rd region's electrodes as seen in Figure 1. The EGG records were obtained from 15 healthy volunteers and 9 gastroparesis patients using electrode placements that indicated in ref. (10). All calculations were performed for two channel EGG signals to determine best electrode placement.

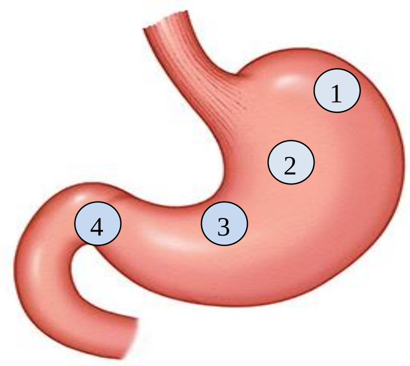


Figure 1. Electrode placements.

2.2. Gastric Emptying

Gastric emptying is the passage of food from pylorus to duodenum. It is provided by the peristaltic motion of the antrum region. Repeated peristaltic movements occur approximately three times cpm in healthy people. The food is divided into smaller pieces with the movement.

Peristaltic movement of the stomach is derived from depolarization of smooth muscle cells in plasma membranes. The oscillations in membrane potential are known as slow waves. Slow waves are generated by the muscles of the proximal corpus along the greater curvature (region of stomach). Greater curvature and region of stomach are shown in Figure 2 [7]. Studies show that slow waves are formed by pacemaker cells. Pacemaker cells are also named as interstitial cells of Cajal (ICCs). Slow waves of the stomach spread with small changes in the stomach. Slow waves cannot be reproduced when there is no ICCs connection, i.e. when the stomach muscles lose their ICCs. For this reason, the stomach cannot be controlled by slow waves, and peristaltic movement of stomach is impaired [13].

Impairment of the stomach peristaltic movements or changes in frequency may cause to disrupt stomach work or delay normal gastric emptying. The rate of gastric emptying varies based on the nervous and hormonal effects of both the stomach and duodenum. Increasing the nutrient volume at the stomach also increases the emptying rate of the stomach. The stomach is usually emptied four hours after eating. The emptying of liquids is faster than solid food. Longer or faster stomach emptying time causes several stomach diseases [14, 15].

Horowitz et al. indicated that solid matters left stomach after the liquids with a delay of 58% and 30% for type 1 and type 2 diabetics [15].

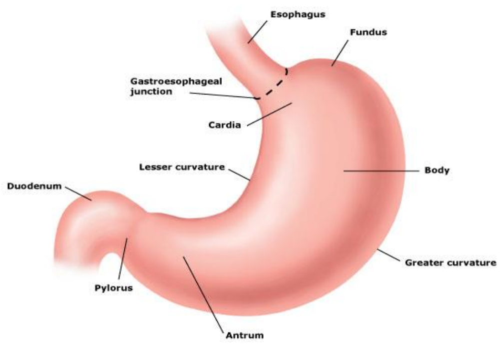


Figure 2. Regions of stomach [16].

Abnormalities in electrical activity of the stomach are called dysrhythmias. There are two types of dysrhytmia, tachygastria and bradygastria. Tachygastria is an increase in the frequency of normal electrical activity of the stomach (3 cpm). Bradygastria is a decrease in the frequency of the slow wave [1-3].

2.3. Gastroparesis

Gastroparesis, which means stomach paralysis, is a gastrointestinal disease caused by diminished or stopped stomach movements. Disease has symptom of delayed gastric emptying without any mechanical obstruction at the out of the stomach. Symptoms of gastroparesis include early feeling of fullness, nausea, vomiting, swelling of the abdomen, epigastric pain and weight loss. These symptoms can vary depending on the person. Different symptoms may not be specific for gastroparesis and symptoms may indicate other diseases. There isn't always an objective relation to the degree of gastric emptying due to the diversity of symptoms. Gastroparesis can occur due to diabetes, neurological, drug utilization or unexplained events. Diabetes is seen in about one-third of patients who are diagnosed with gastroparesis. One of the major complications of diabetes is diabetic gastroparesis [17,18].

Various methods are used for the diagnosis of gastroparesis. Endoscopy, gastric scintigraphy, gastric manometry, EGG, breath test, ultrasonography and intra-gastric imaging with wireless capsule are used for diagnosis. These methods are used to determine delayed gastric emptying [18]. Treatment of gastroparesis is performed with blood sugar control (for diabetic patients), diet, drug therapy, antiemetic therapy, endoscopic therapy with botulinum injection, gastric electrical stimulation and surgical approach [19].

The use of many methods in diagnosis and treatment of gastroparesis and the lack of definite diagnosis have increased gastroparesis as a field of study.

Recently, the gold standard of diagnosis is accepted as gastric scintigraphy. In this method, radioisotope food is given to the subjects to be imaged and the constellation of stomach is examined using gamma camera. Two hours after giving radionuclide, if half of the food is still in the stomach, it is diagnosed as gastroparesis. The duration of the scintigraphy is important and should be long (e.g. 1, 2, or 4 hours). Also, expensive hardware is required for scintigraphy. When these disadvantages are considered, there is a need for a cheap and harmless method of diagnosis of gastroparesis [18,19]. Therefore, in this study EGG is used for diagnosis of gastroparesis.

2.4. Cramer Rao Lower Bound (CRLB)

CRLB is a general method used in statistical signal processing. It is accepted as a statistical method due to the use of probability density function for estimation. CRLB is used to determine the location of the target in radar systems [20], in submarine systems [21], in fingerprinting [22], in Doppler systems [23], and in determination of electrical dipole in EEG / MEG systems [24]. When studies are examined, the CRLB method is mostly used in areas such as target location, radar, submarine, navigation, communication systems, frequency estimation, interval estimation, Autoregressive (AR) parameter estimation and parameter estimation.

Parameter estimation of the system or model is important in analyses of the time series. Different methods are used to examine the accuracy of the estimated parameter. The method used for parameter estimation needs to have the lowest variance to make the closest estimate to the real value. Then, the method is considered to be the best estimation method. The variance bound, called Cramer-rao bound, is determined to estimate performance of the method. The estimate must be unbiased and have low variance so that the estimating process can be performed correctly. θ is parameter vector to be estimated, $\hat{\theta}$ estimated parameter vector, the variance of the estimate is calculated as in Eq. 1 [25].

$$\mathrm{var}(\hat{\theta}) = E\left[\left(\hat{\theta} - E(\hat{\theta})\right)\right] \tag{1}$$

In this equation E [.] is the expected value operator. The optimal estimate is determined by the mean value approach. If the variance is close to zero, the estimate is more accurate [25]. Bias is calculated by the equation $b(\hat{\theta}) = \theta - E(\hat{\theta})$. The bias gives measure of average deviation from actual value. If the estimation is unbiased, it is successful. It must be $E(\hat{\theta}) = \theta$ for unbiased condition. In the CRLB method, the probability density function is used for estimation. The probability density function including the unknown parameter is called *likelihood function* and is shown as $p(x;\theta)$. $\theta = [\theta_1, \theta_2, \theta_3, \ldots, \theta_p]^T$ is the parameter vector to be estimated. In this method, first the natural logarithm of the likelihood function is calculated (In $p(x;\theta)$). Cramer- Rao lower bound is examined as follows [25]:

$$\mathrm{var}(\hat{\theta}_i) \geq \left[I^{-1}(\theta)\right]_{ii} \tag{2}$$

Given $\left[I(\theta)\right]$ in Eq. 2 is fisher information matrix. This matrix is calculated by the equation as in Eq. 3 [25].

$$[I(\theta)]_{ij} = -E\left[\frac{\partial^2 Inp(x;\theta)}{\partial\theta_i \partial\theta_j}\right] \quad i=1,2,\ldots,p \quad ve \quad j=1,2,\ldots,p \tag{3}$$

As can be seen from the above equation, the CRLB calculation is performed by taking the inverse of the Fisher information matrix. The Fisher matrix is obtained by computing the expected value of the derivative of the probability density function [25].

2.5. AR Parameter Estimation with CRLB

Kay obtained the Cramer Rao lower bound for AR parameters using the Power Spectral Density (PSD) of the AR model. The PSD of constituted by the estimated parameters for the {x [0], x [1],..., x [N-1]} data set is as follows [25]:

$$\hat{P}_{xx}(f) = \frac{\hat{\sigma}_u^{\,2}}{\left|1+\sum_{m=1}^{p}\hat{a}[m]e^{-2\pi jfm}\right|^2}$$
$$P_{xx}(f;\theta) = \frac{\hat{\sigma}_u^{\,2}}{\left|A(f)\right|^2} \tag{4}$$

$\theta = [a[1], a[2], \cdots a[p]\sigma_u^2]^T$ are parameters to be estimated. According to CRLB method, first $\frac{\partial InP_{xx}(f;\theta)}{\partial a[k]}$ and $\frac{\partial InP_{xx}(f;\theta)}{\partial \sigma_u^2}$ derivatives are calculated. After calculation of the Fisher information matrix, the lower bound for the AR parameters and the noise variance is calculated as in Eq. 5 [25].

$$\mathrm{var}(\hat{a}[k]) \geq \frac{\sigma_u^{\,2}}{N}[R_{xx}^{\;-1}]_{kk}$$
$$\mathrm{var}(\sigma_u^{\,2}) \geq \frac{2\sigma_u^{\,4}}{N} \tag{5}$$

In Eq. 5 R_{xx} is *pxp* dimensional Teoplitz matrix with $[R_{xx}]_{ij} = r_{xx}[i-j]$ equation.

Several AR parameter estimation methods are used in this study. The CRLB variance lower limit values are obtained for the estimated parameters of patients and healthy subject's signals. The best parameter estimation method was chosen based on this value and PSD calculation was performed using this method.

2.6. Power Spectral Density: Covariance Method

In model-based PSD methods, primarily parameters estimation is done for x[n] (0≤n≤N) data, then spectrum is obtained with these parameters. In the studies, AR parameters are used as a feature for classification patient/healthy subjects, for modeling or noise elimination. There are studies that used AR parameters to differentiate normal and abnormal EGG signals. In this area, a study did by Lin et al. attracted the attention. Lin et al. used Autoregressive Moving Average (ARMA) parameters for classifying EGG signals as normal or abnormal using neural networks [26].

AR models are all zero models. The PSD obtained for the AR (p) model shown in Eq.6 is also explained in Eq.7.

$$y(n) = -\sum_{k=1}^{p} a(k)y(n-k) + x(n) \tag{6}$$

$$P(f) = \frac{\sigma^2}{\left|A(f)\right|^2}$$
$$A(f) = 1 + a_1 e^{-j2\pi f} + \cdots + a_p e^{-j2\pi fp} \tag{7}$$

Autocorrelation method is generally used in AR parameter estimation. The difference of the covariance method is that matrices in calculations aren't Teoplitz matrices. AR parameter calculation for covariance method is given in Eq. 8. In this equation r_x is autocorrelation function. The advantage of the covariance method is that it doesn't apply window to data for the autocorrelation calculation [27].

$$\begin{bmatrix} r_x(1,1) & r_x(2,1) & \cdots & r_x(1,p) \\ \vdots & & & \\ r_x(p,1) & r_x(p,2) & \cdots & r_x(p,p) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = -\begin{bmatrix} r_x(1,0) \\ \vdots \\ r_x(p,0) \end{bmatrix} \tag{8}$$
$$r_x(k,l) = \sum_{n=p}^{N} y(n-l)y^*(n-k)$$

Final Prediction Error (FPE), Akaike Information Criterion (AIC) and Minimizes the Description Length (MDL) methods are used for choice of AR model degree. But in our study we defined model degree according to classification success and CRLB method.

In this study model parameters were used for 4^{th}, 5^{th} and 6^{th} degrees. With these three degrees, features from the PSD function were obtained. Maximum area under the ROC defined highest classification success.

2.7. ROC Curve

The ROC curve is a method aimed at patient/healthy separation. This method is frequently used in medicine. The area under the ROC curve determines the accuracy of patient and healthy discrimination. If this area is 1, then perfect separation is provided. As seen in Figure 3, area is 0.5, interpreted as a worthless test [28].

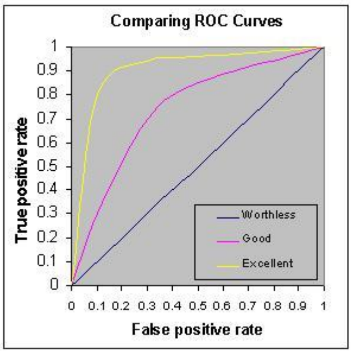


Figure 3. ROC curves according to performance assessment [28].

3. Results

In this study, 4^{th}, 5^{th} and 6^{th} AR parameters were calculated from EGG signals of healthy individuals and patients by using Covariance, Burg, Yule-walker, Least Squares and Modified Covariance methods. To compare the methods, the CRLB variance value of these parameters was obtained. The method with the lowest variance value and the highest value method were selected and AR parameters were estimated by these two methods in all signals. CRLB values for averages of the 4^{th} order AR parameters are shown in Table 1 and Table 2.

When Table 1 and Table 2 are examined, it is seen that covariance method gives the lowest value and Yule-Walker method gives the highest value. For this reason, AR parameters and CRLB were calculated by using covariance and Yule-walker method on EGG signals of all patients and healthy individuals.

Table 1. CRLB values for various analysis methods

Data	Method	CRLB value
Patient on empty stomach	Burg	0.937500181 E-06
	Covariance	0.937500144 E-06
	Modified Covariance	0.937500181 E-06
	Yule-Walker	0.937500229 E-06
	Least-Square	0.937500144 E-06
Healthy on empty stomach	Burg	0.937500195 E-06
	Covariance	0.9375001909 E-06
	Modified Covariance	0.9375001909 E-06
	Yule-Walker	0.937500259 E-06
	Least-Square	0.937500191 E-06

Table 2. CRLB values for various analysis methods

Data	Method	CRLB value
Patient on full stomach	Burg	0.937500154 E-06
	Covariance	0.93750015 E-06
	Modified Covariance	0.937500154 E-06
	Yule-Walker	0.937500194 E-06
	Least-Square	0.93750015 E-06
Healthy on full stomach	Burg	0.937500126 E-06
	Covariance	0.937500128 E-06
	Modified Covariance	0.937500128 E-06
	Yule-Walker	0.937500131 E-06
	Least-Square	0.937500128 E-06

4^{th}, 5^{th} and 6^{th} degrees of AR parameters were estimated by covariance and Yule-walker methods for the EGG signals of patients and healthy people for fasting and satiety stages (for both channels). In Table 3, the mean and standard deviation of 4^{th} degree AR parameters calculated by Covariance and Yule-Walker method were given for first channel data of patients at fasting state.

Table 3. AR parameter estimation

Method	a1	a2	a3	a4
Cov.	-1.41 ∓ 0.23	0.29 ∓ 0.2	0.014∓0.01	0.1∓0.11
Yule-Walker	-1.29 ∓ 0.16	0.15 ∓ 0.1	0.017∓0.01	0.11∓0.11

As in Table 3, the AR parameters were calculated using two methods for three degrees. This method is applied to all signals and the average of the parameters is examined. In other words, the a1 parameter was averaged for 9 patients and the Cramer-Rao lower bound was calculated. Finally, the averages of the CRLB values for the separately calculated parameters a1, a2, a3, a4 were obtained. For instance, Table 4 gives the CRLB values for the first channel of EGG signals obtained by patients.

Table 4. CRLB values of first channel in patients

AR Parameter Estimation Method	4^{th} degree	5^{th} degree	6^{th} degree
Covariance	5.25E-07	6.60E-07	7.53E-07
Yule-Walker	6.32E-06	1.09E-05	1.10E-05

Table 5. CRLB values of second channel in patients

AR Parameter Estimation Method	4^{th} degree	5^{th} degree	6^{th} degree
Covariance	1.36E-07	3.37E-07	4.72E-07
Yule-Walker	5.84E-07	6.93E-07	7.62E-07

Table 6. CRLB values of first channel in healthy people

AR Parameter Estimation Method	4^{th} degree	5^{th} degree	6^{th} degree
Covariance	5.95E-07	7.12E-07	7.93E-07
Yule-Walker	6.85E-07	7.77E-07	8.39E-07

Table 7. CRLB values of second channel in healthy people

AR Parameter Estimation Method	4^{th} degree	5^{th} degree	6^{th} degree
Covariance	5.57E-07	6.77E-07	7.60E-07
Yule-Walker	7.13E-07	7.98E-07	8.55E-07

Table 4-7 shows that the CRLB value is lower in the covariance method. From this result, it was concluded that this method will be more successful to obtain PSD. When the CRLB values for three different degrees were examined, it was seen that the AR parameters at the 4^{th} degree were lower in both methods. For this reason, it shows that obtaining the PSD functions from the 4^{th} degree with the covariance method will give more accurate results. The PSD obtained from a patient's EGG signal is shown in Figure 4.

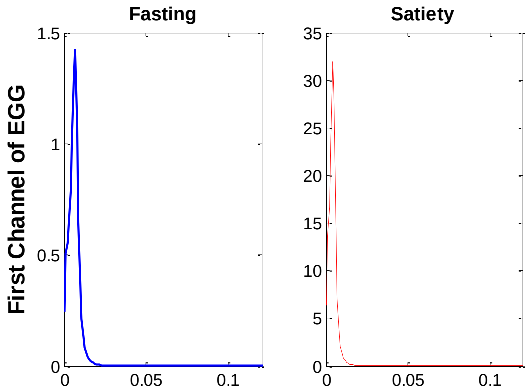


Figure 4. Graphical representation of PSD for a patient

Identification of feature vectors is very important for classification studies. Because attributes that are well chosen and distinctive for the signal enhance the success in classification. For this reason, it is necessary to pay attention to the analysis of the signal and to select those with high separation efficiency from the obtained characteristics. In order to be used in classification, area calculations as features have been carried out numerically on the PSD graphics as shown in Figure 5.

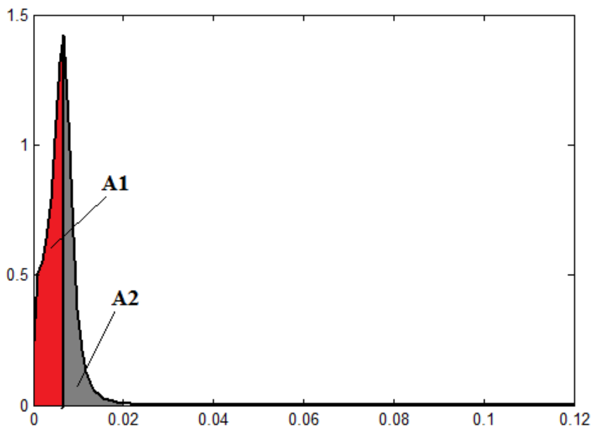


Figure 5. Area account on graphic obtained from PSD function

This field is calculated to the point where the chart is maximized (A1 field is expressed in red color) and after maximum (A2 field is expressed in gray color). Four features extracted from PSD functions, these features:

- A1 area value
- A2 area value
- The ratio of A1 area to total area
- The ratio of A1 area to A2 area

These values separately calculated for first-channel and second channel EGG obtained from patients and health subjects. According to area under the ROC curves results, there is no successful outcome for the first channel in the fasting and satiety stage between the patient and healthy groups. This result shows that EGG signals belonging to patients and healthy people are similar for the first channel in terms of frequency characteristic.

Table 8. Area values under the ROC curve

Second Channel	**A1 Area**	**A2 Area**	**A1/Total**	**A1/A2**
Hungry	0,919	0,978	0,978	0,881
Satiate	0,956	0,844	0,844	0,904

According to the results obtained from second channel EGG signals in Table 8, A1, 91%, A2 area 97.8%, A1/Total area 97.8% and A1/ A2 88.1% can differentiate between patient and healthy group. In addition, in the satiety stage, the A1 area 95.6%, the A2 area 84.4%, A1/Total value 84.4% and A1/A2 90.4% classify the patient and healthy group successfully.

4. Conclusions

When the frequency analysis studies with EGG signals are examined, it is seen that FT, FFT, STFT and Welch method are mostly used. In this study, parametric methods such as Burg, Covariance, Modified Covariance, Least Squares and Yule-Walker were used instead of the nonparametric methods in the literature. The most successful method for estimating AR parameters was decided by the CRLB method. The covariance method gives a lower variance value in AR parameter estimation results obtained from EGG signals from healthy subjects and patients. For this reason, it was decided to use the covariance method for the calculation of the PSD values. The 4^{th}, 5^{th} and 6^{th} degree PSD functions were obtained from the patient and the healthy EGG signals using the covariance method determined by the CRLB method. When results were evaluated by ROC curve, the success rate of 4^{th} degree PSD was 97%, while the success rate of 5^{th} degree PSD was 85% and the success rate of 6^{th} degree PSD was 92%. Thus, for the AR parameters, the choice of 4^{th} degree is more accurate. This proves the validity of the CRLB method. The model degree is determined as the 4^{th} degree giving the smallest CRLB value. Thus, the features that are obtained from PSD functions to identify patients and healthy groups could become more distinctive.

The areas shown in Figure 5 are calculated from the PSD functions generated by the covariance method. The features obtained from first channel data of EGG signals for fasting and satiety stages were not sufficient to differentiate the patient and the healthy group. All the features calculated from the PSD functions for the second channel signals, which were obtained from the patient and healthy people at fasting and satiety stage, differentiate the groups according to the ROC curve results.

When the A1 area is examined for patients and healthy subjects in the stage of fasting, this area value is higher for patients. The high value of this area indicates bradygastria in patients at fasting state. Thus, it was determined that bradygastria occurred for the patients while healthy people were within the normal frequency range according to PSD functions.

As shown under the heading "Data Acquisition" in the second part, the second channel signals correspond to the point at which the gastric pacemaker, i.e., the stomach movement is initiated and advanced. It is shown that according to the second channel data, there are differences in the electrical activity of the stomach between the healthy subjects and patients from EGG signal.

Based on the results of the frequency analysis, it was concluded that the bradygastria were shown for the patients as seen PSD functions, and that the gastric electrical activity was slowed down. In this study, electrode placement and features were obtained which can differentiate EGG signals obtained from healthy and patient subjects. As a result, the proposed methods have obtained the features that will reveal the differences between two groups. It has been concluded that for patients with gastroparesis disease, signals obtained from second channel

(2-3 electrode placement) are more successful to classify patients/healthy groups. Obtained results show that, using CRLB and PSD methods can help diagnose gastroparesis and contribute to the literature.

5. Acknowledgment

In this study, signals are obtained from volunteers by approval of ethics committee of Erciyes University Clinical Research Ethics Committee (approval numbered 2015/115).

6. References

[1] R.M. Stern, K.L. Koch, W.R. Stewart, MW. Vasey, "Electrogastrography: current issues in validation and methodology", *Psychophysiology*, vol. 24, pp. 55–64, 1987.
[2] R.D. Rothstein, A. Alavi, J.C. Reynolds, "Electrogastrography in patients with gastroparesis and effect of longterm cisapride", *Digestive Diseases and Sciences*, vol. 38, pp.1518-1524, 1993.
[3] J.D. Chen, Z. Lin, J. Pan, R.W. McCallum, "Abnormal gastric myoelectrical activity and delayed gastric emptying in patients with symptoms suggestive of gastroparesis", *Digestive Diseases and Sciences*, vol. 41, no. 8, pp. 1538-1545,1996.
[4] W.C. Alvarez, "The electrogastrogram and what it shows", *JAMA*, vol. 78, pp. 1116–1118, 1922.
[5] P. Camborova, P. Hubka, I. Sulkova, I. Hulin, "The pacemaker activity of interstitial cells of Cajal and gastric electrical activity", *Physiological Research*, vol. 52, pp. 275-284, 2003.
[6] B. Bolen. (2014, August). *Stomach Problems Causes*. Available: http://ibs.about.com/od/symptomsofib1/a/StomachProblemsCause.htm.
[7] K.L. Koch and R.M. Stern, "Handbook of Electrogastrography", Oxford Press, New York, 2004.
[8] AJPM. Smout, EJS. Van der, JL. Grashuis, "What is measured in electrogastrography*?", Digestive Diseases and Sciences*, vol. 25, pp. 179-187, 1980.
[9] M. Akay, "Time-frequency analyses of the electrogastrogram", pp. 147-180. *In*: Time Frequency and Wavelets in Biomedical Signal Processing (Eds. M. Akay). The Institute of Electrical and Electronics Engineers, Inc., New York, 1998.
[10] Ç.G. Acer, "Gastroparezi Hastalığının Teşhisi İçin Elektrogastrogram Sinyallerinden Özellik Çıkarımı", M.S. Thesis, Biomed. Eng., Erciyes Uni., Insti. of Sci. and Tech., Kayseri, Turkey, 2015.
[11] Ç.G. Altıntop, F. Latifoğlu, E. Çelikzencir, G. Can Sezgin, M.A. Yurci, "Cramer-Rao Alt Sınırı ile Elektrogastrogram Sinyallerinden Özbağlanım Parametrelerinin Belirlenmesi", Tıp Teknolojileri Ulusal Kongresi, Antalya, Türkiye, 2016, pp.18-21.
[12] T. J. Contreras, M. A. Mayen, A. Servin, L. I. Garay, "System for acquisition and analysis of multichannel electrogastrogram", *7th International Conference on Electrical Engineering, Computing Science and Automatic Control*, Mexico, 2010, pp. 220-224.
[13] T. Ordog,, I. Takayama,, W.K.T. Cheung, S.M. Ward, K.M. Sanders, "Remodeling of networks of interstitial cells of Cajal in a murine model of diabetic gastroparesis", *Diabetes*, vol. 49, pp. 1731-1739, 2000.
[14] E.N. Marieb, K. Hoehn, "Human Anatomy & Physiology", Pearson Higher Education, Cloth, 2006.
[15] M. Horowitz, D. Donovan, K.L. Jones,, C. Feinle, C.K. Rayner, M. Samsom, "Gastric emptying in diabetes: clinical significance and treatment", *Diabetic Medicine*, vol. 19, no. 3, pp. 177-194, 2002.
[16] (2015, May). *Anatomy of Stomach*. Available: http://www.aboutcancer.com/stomach_anatomy.htm
[17] R. Emral, "Diabetic Gastroparesis", *Journal of Ankara Medical School*, vol. 24, no. 3, pp. 129-136, 2002.
[18] H. P. Parkman, R. W. McCallum, "Gastroparesis Pathophysiology, Presentation and Treatment", Humana Press, London, 2012.
[19] M.B. Aljarallah, "Management of diabetic gastroparesis", *Saudi Journal of Gastroenterology*, vol. 17, no. 2, pp. 97-104, 2011.
[20] H. Godrich, A.M. Haimovich, R.S. Blum, "Target localization accuracy gain in MIMO radar-based Systems", *Information Theory, IEEE Transactions on*, vol. 56, no.6, pp. 2783-2803, 2010.
[21] X. Wen, L. Jianlong, "Study of statistical signal models in low-frequency underwater acoustic applications", *Journal of Geophysical Research*, pp. 1-5, 2011.
[22] T. Oktem, D. Slock, "Cramer-rao bounds for power delay profile fingerprinting based positioning", *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, 2011, pp. 2484-2487.
[23] G. Baral-Baron, E. Lahalle, G. Fleury,, X. Lacondemine, J. Schlotterbeck, "Signal parameters estimation using time-frequency representation for Laser Doppler Anemometry", *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*,2012, pp. 2318- 2322.
[24] C.H. Muravchik, A. Nehorai, "EEG/MEC error bounds for a static dipole source with a realistic head model", *Signal Processing, IEEE Transactions on*, vol. 49, no. 3, pp. 470-484, 2001.
[25] S.M. Kay, "Fundamentals of Statistical Signal Processing Estimation Theory", Prentice-Hall, Englewood Cliffs, NJ, 1993.
[26] Z. Lin, J. Chen, "Classification of the normal and abnormal electrogastrogram using back-propagation neural Networks", *Engineering in Medicine and Biology Society, Engineering Advances: New Opportunities for Biomedical Engineers. Proceedings of the 16th Annual International Conference of the IEEE*,pp. 1091-1092, 1994.
[27] J.G. Proakis, D.G. Manolakis, "Digital Signal Processing Principles, Algorithms and Applications". Prentice Hall: New Jersey, 1996.
[28] Thomas G. Tape. (2016, December). *Interpreting Diagnostic Tests.* Available: http://gim.unmc.edu/dxtests/roc3.htm

Effects of Autapse and Channel Blockage on Firing Regularity in a Biological Neuronal Network

Rukiye UZUN[1], Mahmut OZER [1]

[1]Bulent Ecevit University, Zonguldak, Turkey
rukiye.uzun@beun.edu.tr, mahmutozer2002@yahoo.com

Abstract: *In this paper; the effects of autapse (a kind of self-synapse formed between the axon of the soma of a neuron and its own dendrites) and ion channel blockage on the firing regularity of a biological small-world neuronal network, consists of stochastic Hodgkin-Huxley neurons, are studied. In this study, it is assumed that all of the neurons on the network have a chemical autapse and a constant membrane area. Obtained results indicate that there are different effects of channel blockage and parameters of the autapse on the regularity of the network, thus on the temporal coherence of the network. It is found that the firing regularity of the network is decreased with the sodium channel blockage while increased with potassium channel blockage. Besides, it is determined that regularity of the network augments with the conductance of the autapse.*

Keywords: *Autapse, channel blocking, ion channel noise, small-world neuronal netwoks.*

1. Introduction

The nervous system, having a complex biologic structure, comprises of billions of nerve cells (neurons) and synaptic connections between them [1]. In this complex structure, it is believed that the signal communication (i.e. information transmission) can be carried out by synaptic coupling [2, 3]. Synapses are essentially categorized in two different types: electrical and chemical synapses [4]. However, several decades ago, neurobiologists found out that some neurons could be connected to itself which forms a time-delayed feedback mechanism on a cellular level [5-9]. These self-synapses named as an autapse and were proposed by Van der Loss and Glaser in 1972 [10]. Although autapses have an unusual (odd) structure, they have been observed commonly in various brain areas, such as neocortex, hippocampus and cerebellum etc. [4, 11-13, Wang and Chen 2015, Wang et al 2016].

Besides, it has been experimentally presented that these type of synapses have significant effects on the firing dynamics of neurons. For example, Bacci and Huguenard (2006) showed that the autaptic connections have a vital role on determining the precise spike timing of inhibitory interneurons in neocortex [14]. Rusin et al. [15] demonstrated that usage of continuous time-delayed feedback stimulation can arrange the synchronization of spikes in cultured neurons. In addition to these experimental studies, there is also numerous studies that investigate the impacts of autapse on firing dynamics of the neuron with computational neuronal models [16-25]. In this context, Li et al [16] investigated the effect of ion channel noise on dynamics of neuron in the presence of autapse based on a stochastic Hodgkin-Huxley (HH) model. They found that autapse is reduced the spontaneous spiking activity of neuron at characteristic frequencies by inducing bursting firing and multimodal interspike interval distribution. Connelly [17] revealed that autapses augment the gamma oscillations of interneurons, known as basket cells, in a Wang-Buzsáki model during the gamma oscillations. Wang et al [18] discovered that autapses causes to switch the dynamics of electrical activity of Hindmarsh-Rose model among different firing patterns (quiescent, periodic and chaotic). In another study, it is demonstrated that autapses can regulate the mode-locking behaviors of a HH model subjected to sinusoidal stimulus, especially by autaptic delay time [19]. Recently, Yılmaz and Ozer [20] consider the weak signal detection performance of a stochastic HH model with regard to an electrical autapse. They obtained that the detection performance can be healed or deteriorated depending on autapses' parameters. Similar studies concerning weak signal detection performance have been realized on small-world (SW) and scale-free (SF) neural networks by Yilmaz et al and have been obtained analogous results [21, 22]. Wang and Gong [23] have shown that both multiple coherence resonance (MCR) and synchronization transitions are occurred because of autpases, in SW neuronal network. Recently, Wang and Chen revealed that autapses provide an opportunity to engineer the response of a HH model to subthreshold stimulus [24]. Moreover, in this study, it is revealed out that the detection of subthreshold stimulus is increased due to autapse. More recently, the impacts of both electrical and chemical autapses on the temporal coherence

Received on: 30.01.2017
Accepted on: 14.03.2017

of a single HH model and SF neuronal network is analyzed by Yilmaz et al [25]. They obtained that the dynamics of neuron or network change prominently with the proper choice of autaptic parameters.

Nonetheless, it is well known that neuronal-information processing operates under diverse noise sources, which has therapeutic influences contrary to expectations [26]. The most important noise source in neural systems is originated in voltage-gated ion channels due to their random transitions between conducting and non-conducting states [27]. The strength of channel noise is related with the numbers of ion channels on membrane, thus membrane area. But, its real effect on neuron's dynamics is determined by the number of active ion channels participating in the generation of spikes [28]. Therefore, addressing impacts of the number of active ion channels is important especially to uncover the role of specific ion channel noise on neuronal spiking activity. Experimental studies indicate that some neurotoxins such as tetraetylammonium (TEA) ve tetratoxin (TTX) can alter the properties of ion channels [29]. For a given membrane, by a fine-tuned addition of these toxins a certain portion of potassium- and sodium ion channels could be disabled or blocked and hence the number of active (working) ion channels can be reduced. There are also plenty of studies based on different computational neuron models, where the impacts of changing the number of particular ion channels on firing dynamics of a single neuron or neuronal networks' is examined [30-38]. In these studies, it is uncovered that channel blocking has crucial impacts on firing dynamics.

In this study, different from the above studies, we investigate how the firing dynamics of Newman-Watts SW neuronal network, consisting of stochastic HH neurons, change in the presence of both ion channel blocking and autapse. For this aim, we assume that all neurons on the network have a chemical autapse and are exposed to same ratio of ion channel blocking.

2. Model and Method

Time evolution of membrane potentials of Newman-Watss network of HH neurons in the presence of chemical autapse is given as follows [16, 21-23, 25]:

$$C_m \frac{dV_i}{dt} = -G_{Na}(m_i,h_i)(V_i - V_{Na}) - G_K(n_i)(V_i - V_K) - G_L(V_i - V_L) + \sum \varepsilon_{ij}\left(V_j(t) - V_i(t)\right) + I_{aut_i} \quad (1)$$

Here C_m is the membrane capacitance per unit area, V_i denotes the membrane potential of neuron $i=1,...,N$ (N shows the network size) in mV, and $V_{Na}=50mV$, $V_K=-77mV$ and $V_L=54.4mV$ are the equilibrium potential of sodium, potassium and leakage ion channels, respectively. ε_{ij} denotes the coupling strength between neurons i and j, whereby we set $\varepsilon_{ij}=0.1$ if there is a connection between neurons or $\varepsilon_{ij}=0$ otherwise.

I_{auti} is the delayed feedback current related to the autaptic connection of neuron i. The autaptic current of a simplified neuron model can be described by two different forms. One is linear coupling (i.e. electrical) and the other is nonlinear coupling (i.e. chemical). In this study we use chemical autaptic current which is given by the following two equations [21,23,41]:

$$I_{aut_i} = -g_{aut}\left[V_i(t) - V_{syn}\right]S(t-\tau); \quad (2a)$$
$$S_i(t-\tau) = 1/\left\{1 + \exp[-k(V_i(t-\tau) - \theta]\right\} \quad (2b)$$

Here g_{aut} is the autaptic intensity and τ is the delay time which is taken as 10 ms throughout the study. V_{syn} represents the autaptic reversal potential and its value varies depending on whether the autapse is excitatory ($V_{syn}=2$ mV) or inhibitory ($V_{syn}=-2$ mV). We assumed that all neurons have excitatory chemical autapse with the same autaptic parameters. The other parameters' values, in equation (2), is set as $k=8$, $\theta=-0.25$. More interpretations of all parameters in Eq. 2 can be found at Ref. [41].

In equation (1), G_{Na}, G_K and G_L denote sodium, potassium and leakage ion channel conductance, respectively. In the model, the leakage conductance is constant, $G_L=0.3mScm^{-2}$, whereas the others dynamically change as follows [30, 34, 37, 38]:

$$G_{Na}(m_i,h_i) = g_{Na}^{max}\chi_{Na}m_i^3h_i,\ G_K(n_i) = g_K^{max}\chi_K n_i^4 \quad (3)$$

where $g_{Na}^{max}=120mScm^{-2}$ and $g_K^{max}=36mScm^{-2}$ is the maximal conductance of the corresponding ion channel. χ_i $i=(Na, K)$ is a scaling factor, which equals to the ratio of active (non-blocked) i-type ion channel numbers to total i-type ion channel number within a given membrane size [30, 31]. This factor is confined to unit interval and i-type ion channel blocking level gets higher as the factor approaches zero. In equation (3); n_i, m_i and h_i are dimensionless variables that represent for the potassium ion channel activation, sodium ion channel activation and inactivation, respectively. The dynamics of these variables is described by the following Langevin equation for a finite membrane size [42]:

$$\frac{dx_i}{dt} = \alpha_{x_i}(V)(1-x_i) - \beta_{x_i}(V)x_i + \xi_{x_i}(t),$$
$$x_i = m,n,h \quad (4)$$

Here α_{xi} and β_{xi} are the voltage-dependent opening and closing rates of x_i. $\zeta_{xi}(t)$ is the zero mean Gaussian white noises whose auto-correlation functions given as follows [42]:

$$\langle \xi_x(t)\xi_x(t')\rangle = \frac{2\alpha_x\beta_x}{N_j\chi_j(\alpha_x+\beta_x)}\delta(t-t') \quad (5a)$$
$$N_j = \rho_j S;\ \ j = Na, K; x = m,n,h \quad (5b)$$

where S denotes the membrane size, N_j and ρ_j are the total number and density of corresponding ion channel, respectively [30, 31]. Let $\rho_{Na}=60\mu m^{-2}$ and $\rho_K=18\mu m^{-2}$. In the model, it is assumed that sodium and potassium ion channel densities are homogenous.

The considered neuronal network is Newman–Watts (NW) small world network [39]. The network is constituted as follows. It starts with a regular ring consisting of $N = 60$ identical HH neurons in which each neuron has connections with its $k=2$ nearest neighbors. Then new links are randomly added between non- adjacent neighbors. The randomness expressed with the fraction of shortcuts corresponding to *p=M/(N(N-1))* (*M* is the number of new links). The constructed network exhibits small-world properties when p is confined to unit interval. In this study, we construct the network with *p=0.15*.

In order to analyze the impacts of ion channel blockage and autapses on dynamics of network, we calculate the collective firing regularity of the network. To do so, we first obtain the firing times of neurons on the network via the average membrane potential () and then calculate the spontaneous collective spiking regularity with:

$$\lambda = \frac{\langle ISI \rangle}{\sqrt{\langle\langle ISI^2 \rangle\rangle - \langle ISI \rangle^2}} \qquad (6)$$

Here $\langle ISI \rangle$ and $\langle ISI^2 \rangle$ are the mean and mean-squared interspike intervals, respectively. The firing pattern of network gets more regular, as λ increases. In the following results, spike times are determined by $V_{avg}(t)$ crosses the detection threshold 0 from below and also each λ value is calculated by averaging over 20 different network realization.

3. Results and Discussion

In a recent study, Wang and Gong [23] have investigated how the temporal coherence (namely, firing regularity) and synchronization of Newman-Watts SW neural network alter in the presence of autaptic self-delayed feedback. In the light of their study, here, we systemically investigate the effects of potassium and sodium ion channel blocking on the collective firing regularity of network (λ) as a function of autaptic intensity (g_{aut}) at a constant autaptic time delay ($\tau = 10ms$). To show effects of each specific ion channels, the density of one channel type is varied while keeping the other channel type equals one ($\chi_K = 1$ or $\chi_{Na} = 1$). Besides, throughout this study, we set $S = 6\,\mu m^2$ because of the optimum temporal coherence is obtained [32, 33]. Then, we calculate λ with respect to g_{aut} for various χ_K and χ_{Na}.

We first investigate how λ changes with the non-blocked potassium ion channel fraction (i.e. active potassium ion channel, χ_K) as a function of g_{aut}, as shown in Figure 1. From the obtained results, one can see that the reduction of the potassium ion channels has two essential effect on λ. First is that λ increases with the increment of the potassium ion channel blocking (i.e. decrease in χ_K) for a constant g_{aut}. Later, for a constant χ_K, λ displays an exponential raise beyond a certain g_{aut}. This value of g_{aut} moves to smaller values as χ_K decreases.

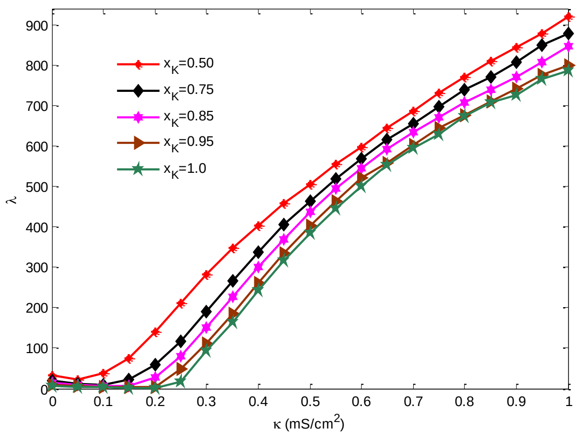


Figure 1. Dependence of collective firing regularity λ on autaptic intensity g_{aut} for various potassium channel block ratios χ_K at a constant autaptic time delay $\tau = 10ms$.($\chi_{Na} = 1.0,\ p = 0.15, \varepsilon = 0.1, S = 6\ \mu m^2$).

To gain more insight into the dependence of collective firing regularity on χ_K, we further analyze the variation of λ as a function of χ_K. We set autaptic parameters as $\tau = 10ms$ and $g_{aut} = 0.15\ mScm^{-2}$. Obtained result is presented in Figure 2. It can be obviously seen that λ increases as χ_K decreases. The potassium ion channel blocking causes the increment at the excitability of neurons in the network, and because of this, collective firing regularity of network becomes more regular. Besides, the presence of autaptic connections in the network and the increment of autaptic intensity g_{aut} augments the collective firing dynamics of the network. Briefly, one can conclude that the presence of both potassium ion channel blocking and autaptic connections have therapeutic effects on the temporal coherence of network.

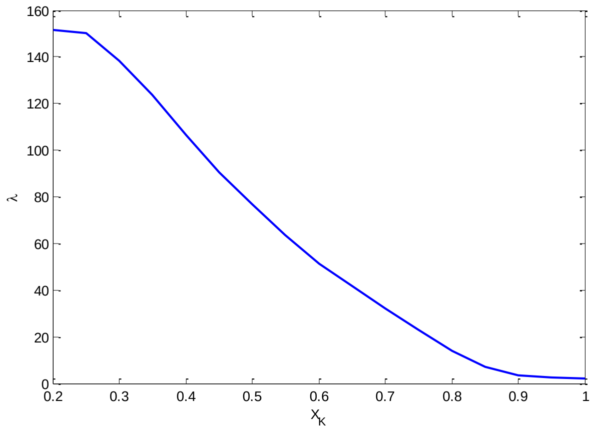


Figure 2. Dependence of collective firing regularity λ on potassium channel block ratios χ_K, when the autaptic parameters are constant ($\tau = 10ms, g_{aut} = 0.15\ mScm^{-2}$). ($\chi_{Na} = 1.0,\ p = 0.15, \varepsilon = 0.1, S = 6\ \mu m^2$).

After determining the impacts of potassium ion channels on temporal coherence of network in the presence of autapse, we examine the effects of sodium ion channels known as another important ion channels at determining the firing dynamics of neuron. For this, we implement similar simulations for sodium ion channels as we did in case of potassium channel block. In the simulations we consider $\chi_{\mathrm{Na}} > 0.7$, because the average membrane potential V_{avg} do not include spikes under below this χ_{Na} value. In Figure 3, it can be clearly seen that reduction in active sodium ion channels influences the temporal coherence of network (λ) oppositely compared to the case of potassium ion channel blockage. Namely, decreasing the χ_{Na} disrupts the λ. But this destructive influence vanishes at smaller g_{aut} and besides, the value of g_{aut} where the sodium ion channel blockage does not make important variation on λ increases as χ_{Na} decreases. As a result, the sodium ion channel block degrades the temporal coherence of the network in the presence of autaptic connections.

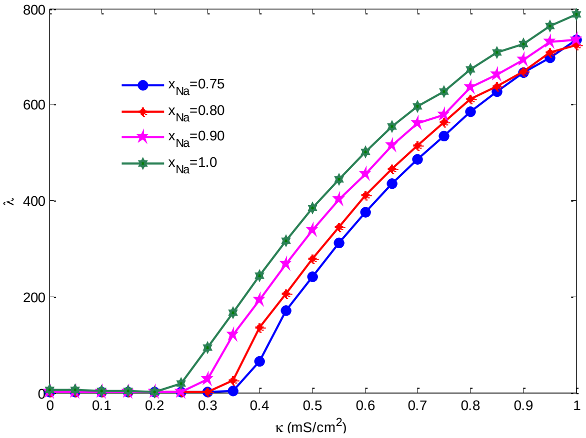


Figure 3. Dependence of collective firing regularity λ on autaptic intensity g_{aut} for various sodium channel block ratios χ_{Na} at a constant autaptic time delay $\tau = 10ms$. ($\chi_K = 1.0$, $p = 0.15, \varepsilon = 0.1, S = 6\ \mu m^2$).

In order to show prominently the impact of sodium ion channel blockage on temporal coherence of the network, we also give the simulation results how λ changes as a function of χ_{Na} at constant autapse parameters ($\tau = 10ms$, $g_{aut} = 0.5\ mScm^{-2}$) (Figure 4). From Figure 4, one can see clearly the destructive impact of the reduction of active sodium ion channel.

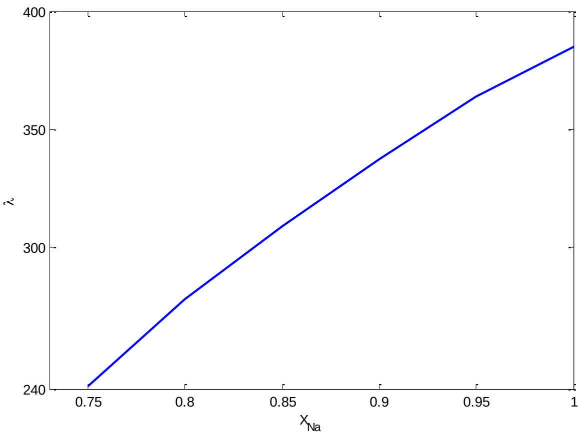


Figure 4. Dependence of collective firing regularity λ on sodium channel block ratios χ_{Na} when the autaptic parameters are constant ($\tau = 10ms$, $g_{aut} = 0.5\ mScm^{-2}$). ($\chi_{\mathrm{K}} = 1.0$, $p = 0.15, \varepsilon = 0.1, S = 6\ \mu m^2$).

5. Conclusions

In sum, we investigate the collective firing regularity of NW small-world neuronal networks in the presence of both autapse and ion channel block. Comparing the effect of potassium and of sodium channel block, one can see that they have much different influences on temporal coherence of the network. For potassium ion channel block, firing regularity increases, while for sodium channel block, it affects negatively. This could be because of the discriminating impacts of potassium and sodium ion channel blockage on the firing dynamics [30, 31].

6. References

[1] R. Uzun, "Biyolojik Nöral Ağlarda Latans Dinamiklerinin Analizi", PhD thesis, Elec Depar. of EEE, BEU, Zonguldak., Turkey, 2014.
[2] E. R. Kandel, J. H. Schwartz, T. M. Jessel, S. A. Siegelbaum, and A. J. Hudspeth, "Principles of Neural Science 4th Edition ", McGraw-Hill Medical, New York, England, 2000.
[3] H. Qin, J. Ma, C. Wang, and Y. Wu, "Autapse-induced sprial wave in network of neurons under noise", PlosOne, vol. 9, no. 6, pp. 1-9, 2014.
[4] E. P. Furshpan, and D. D. Potter, "Transmission at the giant motor synapses of the crayfish", J. Physiol., vol.145, pp. 289-345, 1959.
[5] A. B. Karabelas, and D. P. Purpura, "Evidence for autapse in the substantianigra", Brain Res., vol. 200, no. 2, pp. 467–473, 1980.
[6] G. Támas, E. H. Buhl, and P. Somogyi, "Massive autaptic self-innervation of GABAergic neurons in cat visual cortex", J. Neurosci., vol. 17, no. 16, pp. 6352-64, 1997.
[7] J. M. Bekkers, "Neurophysiology: are autapses prodigal synapses?", Curr. Biol., vol. 8, no. 2, pp. R52-55, 1998.
[8] J. M. Bekkers, "Synaptic transmission: functional autapses in the cortex", Curr. Biol., vol. 13, no. 11, pp. R433-435, 2003.
[9] K. Ikeda, and J. M. Bekkers, "Autapses", Curr. Biol., vol. 16, no. 9, pp. R308, 2006.
[10] H. Van der Loos, and E. M. Glaser, "Autapses in neocortex cerebri: synapses between a pyramidal cell's axon and its own dendrites", Brain Res., vol. 48, pp. 355-360, 1972.
[11] J. Lübke, H. Markram, M. Frotscher, and B. Sakmann, "Frequency and dendritic distribution of autapses established by layer 5 pyramidal neurons in the developing rat neocortex: Comparison with synaptic innervation of adjacent neurons of the same class", J. Neurosci., vol. 16, pp. 3209-3218, 1996.
[12] M. H. Flight, "Neuromodulation: Exerting self-control for persistence", Nat. Rev. Neurosci., vol. 10, pp. 316, 2009.
[13] T. Branco, and K. Staras, "The probability of neurotransmitter release: Variability and feedback control at single synapses", Nat. Rev. Neurosci., vol. 10, pp. 373-383, 2009.
[14] A. Bacci, and J. R. Huguenard, "Enhancement of spike-timing precision by autaptic transmission in neocortical inhibitory interneurons", Neuron, vol. 49, pp. 119-130, 2006.
[15] C. G. Rusin, S. E. Johnson, J. Kapur, and J. L. Hudson, "Engineering the synchronization of neuron action potentials using global time-delayed feedback stimulation", Phys. Rev. E, vol. 84, pp. 066202, 2011.
[16] Y. Li, G. Schmid, P. Hänggi, and L. Schimansky-Geier, "Spontaneous spiking in an autaptic Hodgkin–Huxley setup", Phys. Rev. E, vol. 82, pp. 061907, 2010.
[17] W. M. Connelly, "Autaptic connections and synaptic depression constrain and promote gamma oscillations", PlosOne, vol. 9, pp. e89995, 2014.

[18] H. Wang, J. Ma, Y. Chen, and Y. Chen, "Effect of an autapse on the firing pattern transition in a bursting neuron", Commun. Nonlinear Sci. Numer. Simul., vol. 19, pp. 3242-3254, 2014.
[19] H. Wang, Y. Sun, Y. Li, and Y. Chen, "Influence of autaptic self-feedback on mode-locking structure of a Hodgkin–Huxley neuron under sinusoidal stimulus", J. Theoret. Biol., vol. 358, pp. 25-30, 2014.
[20] E. Yilmaz, and M. Ozer, "Delayed feedback and detection of weak periodic signals in a stochastic Hodgkin–Huxley neuron", Physica A, vol. 421, pp. 455-462, 2015.
[21] E. Yilmaz, V. Baysal, M. Ozer, and M. Perc, "Autaptic pacemaker mediated propagation of weak rhythmic activity across small-world neuronal networks", Physica A, vol. 444, pp. 538-546, 2016.
[22] E. Yilmaz, V. Baysal, M. Perc, and M. Ozer, "Enhancement of pacemaker induced stochastic resonance by an autapse in a scale-free neuronal networks", Sci. China Technol. Sc., vol. 59, no. 3, pp. 364-370, 2016.
[23] Q. Wang, and Y. Gong, "Multiple coherence resonance and sychronization transitions induced by autaptic delay in Newman-Watts neuron networks", Appl. Math. Model., vol. 40, pp. 7147-7155, 2016.
[24] H. Wang, and Y. Chen, "Response of autaptic Hodgkin–Huxley neuron with noise to subthreshold sinusoidal signals", Physica A, vol. 462, pp. 321-329, 2016.
[25] E. Yilmaz, M. Ozer, V. Baysal, M. Perc, and M. Perc, "Autapse-induced multiple coherence resonance in single neurons and neuronal networks", Sci. Rep., vol. 6, pp. 30914, 2016.
[26] P. Hänggi, "Stochastic resonance in biology", Chem. Phys. Chem., vol. 3, pp. 285-290, 2002.
[27] P. N. Steinmetz, A. Manwani, C. Koch, M. London, and I. Segev, "Subthreshold voltage noise due to channel fluctuations in active neuronal membranes", J. Comput. Neurosci., vol. 9, pp. 133-148, 2000.
[28] E. Schneidman, B. Freedman, and I. Segev, "Ion channel stochasticity may be critical in determining the reliability and precision of spike timing", Neural Comput., vol. 10, pp. 1679-1694, 1998.
[29] B. Hille, "Ionic Channels of Excitable Membranes", Unv. of Washington, Sinauer Press, USA, 1992.
[30] G. Schmid, I. Goychuk, and P. Hänggi, "Effect of channel block on the spiking activity of excitable membranes in a stochastic Hodgkin-Huxley model", Phys. Biol., vol. 1, no. 1-2, pp. 61-66, 2004.
[31] G. Schmid, I. Goychuk, and P. Hänggi, "Controlling the spiking activity in excitable membranes via poisoning", Physica A, vol. 344, no. 3-4, pp. 665-670, 2004.
[32] Y. B. Gong, B. Xu, X. G. Ma, and J. Q. Han, "Effect of channel block on the collective spiking activity of coupled stochastic Hodgkin-Huxley neurons", Sci. China Ser. B, vol. 51, no. 4, pp. 341-346, 2008.
[33] M. Ozer, M. Perc, and M. Uzuntarla, "Controlling the spontaneous spiking regularity via channel blocking on Newman-Watts networks of Hodgkin-Huxley neurons", Europhys. Lett., vol. 86, no. 4, pp. 40008-14, 2008.
[34] M. Uzuntarla, R. Uzun, E. Yilmaz, M. Ozer, and M. Perc, "Noise-delayed decay in the response of a scale-free neuronal network", Chaos Soliton Fract., vol. 56, pp. 202-208, 2013.
[35] J. Ma, H. Long, C. N. Wang, and Z. S. Pu, "Robustness, death of spiral wave in the network of neurons under partial ion channel block", Commun. Theor. Phys., vol. 59, no. 2, pp. 233-242, 2013.
[36] L. Huang, J. Ma, J. Tang, and F. Li, "Transition of ordered waves in neuronal network induced by diffusive poisoning of ion channels", J. Biol. Syst., vol. 21, no. 01, pp. 1350002, 2013.
[37] R. Uzun, M. Ozer, and M. Perc, "Can scale-freeness offset delayed signal detection in neuronal networks?", Eur. Phys. J. A, vol. 105, no. 6, pp. 60002, 2014.
[38] R. Uzun, and M. Ozer, "The effect of channel blocking on first spike timing", in IEEE 23nd Signal Processing and Communications Applications Conference, Malatya, Turkey, 2015, pp. 1191-1199.
[39] M. E. J. Newman, and D. J. Watts, "Scaling and percolation in the small-world network model", Phys. Rev. E, vol. 60, no. 1, pp. 7332-7342, 1999.
[40] A. Hodgkin, and A. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve", J. Physiol., vol. 117, no. 4, pp. 500-544, 1952.
[41] N. Burić, K. Todorović, and N. Vasović, "Synchronization of bursting neurons with delayed chemical synapses", Phys. Rev. A, vol. 78, no. 3, pp. 036211, 2008.
[42] R. F. Fox, "Stochastic versions of the Hodgkin–Huxley equations", Biophys. J., vol. 72, no. 5, pp. 2068-2074, 1997.

Dose and Time Controlled Electronic Insulin Pen: Bilensulin

Mana SEZDI[1], Ahmet BILEN[2]

[1]Biomedical Device Technology Programme, Istanbul University, Istanbul, Turkey
[2]Biomedical Engineering, Yeni Yuzyil University, Istanbul, Turkey
mana@istanbul.edu.tr, ahmetbilen_07@hotmail.com

Abstract: ***In this study, an electronic insulin pen was designed for diabetic patients dependent on insulin use in order to prevent wrong timing and wrong dosage caused from old age, vision or mental illness. During design, the patients' necessities such as easy-to-use, have been considered. The insulin pen that can be used electronically for routine dose injections, and that is called as Bilensulin, was designed to be able to used manually in the emergency. It is aimed that the patient can set the dose easily by using the fairly simple menu of the device. Additionally, the device was designed as reusable by changing the insulin cartridge. Because the tube chamber was designed to be suitable for all types of insulin, it can be used by changing cartridge without using different pen. Dose accuracy was tested by performing insulin dose measurements on a sensitive balance, and dose of Bilensulin was compared with the three types of insulin pens. The patent has been applied for this designed electronic insulin pen. The mechanism of change of cartridge has been put into the foreground in terms of ease of use and constituted the sub-structure of the patent application. Currently, the studies on the device are continuing.***

Keywords: *Insulin, dose, diabetes, insulin pen.*

1. Introduction

Insulin is a hormone secreted by the pancreas that helps the body use sugar and keeps the sugar level at the normal level. The sugar, which is the main food source of our bodies, must enter the body cells of the bloodstream (muscle cells, fat cells and liver cells) in order to provide energy. Insulin separates the blood sugar from the bloodstream and allows it to enter the cell, thus reducing the level of sugar in the blood [1].

In a non-diabetic person, after each food intake, the pancreas produces insulin to make the energy of the food taken up. This means that all people are dependent on insulin. In diabetics, the pancreas does not produce enough insulin or the insulin produced is not used by the target cells (muscle, fat and liver cells). In this case, insulin, which has a vital prescription for our bodies, needs to be taken from the outside [1].

Due to the complexity of insulin injections, insulin pens are preferred by the patients with diabetes. Several studies have demonstrated benefits of insulin pen devices for patients especially for pediatric and geriatric patients [2-4]. The studies have shown that insulin pens are preferred because of their ease of use and flexibility. Insulin pens provide more accuracy, especially for low-dose administration and the old patients [5-7].

During insulin therapy, there are some barriers presented from diabet disease in elderly patients.

Visual impairment, blindness, limited joint mobility and symptomatic peripheral neuropathy cause the difficulties on insulin theraphy [8]. These disabilities may make it more difficult for patients to use insulin pen. Visibility required to adjust a required dose and force required to inject an insulin dose are important factors for insulin pens [9]. A device with a clear dose scale, audible clicks accompanying the dialing of each dose, a large dose delivery button, and comfortable to handle device are important features for old patients [8].

There are two types of insulin pens: prefilled disposable pens and refillable pens [10]. Prefilled disposable pens are designed with a built-in and prefilled insulin reservoir containing 3ml (300 units) of insulin. Once empty, the patient must use another prefilled device. Refillable pens combine the reusable syringe and insulin container with disposable insulin cartridges containing 3ml insulin (Figure 1). These pens may be used for several years with only the needles and cartridges being replaced [5].

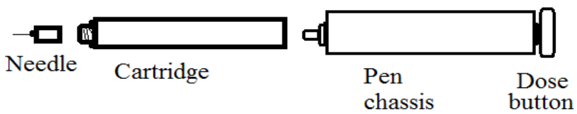


Figure 1. Schematic representation of refillable insulin pens

Before use, the refillable pens must be loaded with an insulin cartridge. The holder of cartridge is twisted from the

Received on: 31.01.2017
Accepted on: 14.03.2017

barrel and the cartridge is loaded with the piston rod fully depressed.

The aspects of the insulin pens for elderly patients must be considered as physical aspects and memory aspects separately [11].

The most important aspects of the physical feature of the insulin pen are attaching/removing cartrige holder and applying low force to inject insulin [12, 13]. But, because diseases such as hand tremor or difficulty in applying adequate force during injection affect the treatment process of elderly patients. In order to remove this negative effect, electronic insulin pens with motor system have begun to be developed.

Whether the insulin pen is mechanical or electronic, the patients and healthcare professionals want to be sure injected dosage and injection time. The insulin pens with memory function developed for this purpose are still used in the market. The feature is designed for patients who cannot remember if they administered the last dose or how much was last injected [14]. The most important aspects of the memory feature of the insulin pen according to the patients and healthcare professionals are [12];

Ability to view the units of previous insulin dose
Ability to view the time of previous insulin dose
Ability to view the date of previous insulin dose
Ability to view the previous 16 doses
Ability to confirm injection taken

In this study, considering the physical and memory features of such insulin systems, an easy-to-use electronic insulin pen was designed to prevent the wrong dose from being used in patients with mental or visual disturbances.

In this study, considering the physical and memory features of such insulin systems, an easy-to-use electronic insulin pen, named as Bilensulin (BS), was designed to prevent the wrong dose for patients with mental or visual disturbances or hand immobilities.

2. Design Modules

Bilensulin (BS) is a reusable pen for use with prefilled insulin cartridges. It is the pen that possesses a memory feature permitting the previous 16 doses to be stored with the date and time of dosing. BS designed for electronic injection, as shown in Figure 2 consists of the components listed below:

1) Motor: The self-driven driver allows the piston to move by turning. It begins to rotate with the force given to it and pushes the piston. When the cartridge is finished, it moves in the opposite direction and pulls back the piston.

2) Piston Driver: It is connected to the motor and has the cartridge on the other end. The motor moves in response to the push and applies pressure to the cartridge to inject the liquid. It is female and has a male part. It performs the movement in both directions according to the screw step.

3) Support Module: During the cartridge exchange, it prevents the piston from coming out of the moving part. It keeps the system fixed.

4) Insulin Cartridge: It is a tube with insulin liquid inside. It contains 3ml (300U) insulin.

5) Cartridge Change Mechanism: This is the part where the device doubles for replacement with the cartridge finish. It looks like a standard door hinge. The driver folds in two as the piston comes out of the cartridge and the cartridge is replaced.

6) Fixer: Half-moon constructions for fixing motor, piston driver and cartridge. It keeps the structures on top of each other and keeps them together during the operation of the device.

7) Needle: The part where insulin in the cartridge is injected into the patient.

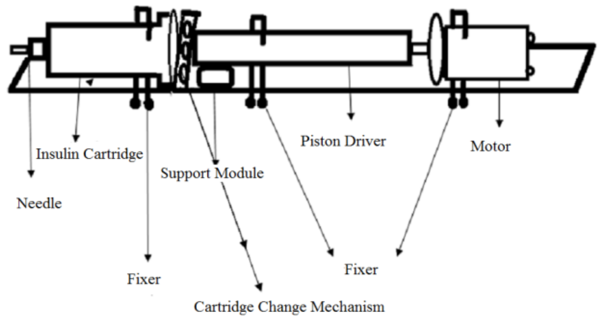


Figure 2. Cartridge change mechanism

The Arduino uno microcontroller was used for circuit design (Figure 3). Arduino is a very popular and easy to use programmable board. It consists of a simple hardware platform and a free source code editor. The Arduino board provides four basic functional elements: An Atmel ATmega328P AVR microcontroller, a simple 5 V power supply, a USB-to serial converter for loading new programs onto the board, I/O headers for connecting sensors, actuators and expansion boards.

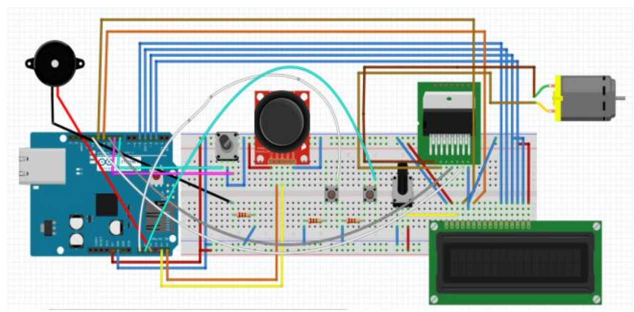


Figure 3. Circuit design of BS

In the programming used when creating the prototype, the software is based on the time taken to complete a round. The time of a tour of BS has been measured, and the software has been completed by using this account. However, in the process of converting the prototype to the final product, it will be more useful to control the device according to whether the motor is turning or not, by using an encoder. Thus, depending on the technical problems that might occur in the motor, it can be controlled whether the motor is turning or not. After all, injection of BS can be also clearly controlled by controlling motor.

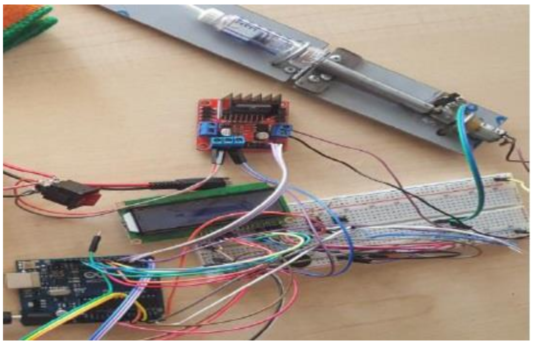


Figure 4. Bilensulin prototype

The features of Bilensulin whose prototype is given in Figure 4 can be explained as: The insulin dose which must be routinely injected by the patient during the day is entered as data at the first opening of the device. It is ensured that the insulin is automatically injected into the patient on time. First, second and third insulin dose adjusted can be followed from LCD monitor (Figure 5). The dose and time value initially entered are fixed until the next change time. In the event of a crisis, the device can be manually converted to the desired dose and manually injected. In overdose situations, the alarm of "emergency" is seen in the LCD monitor.

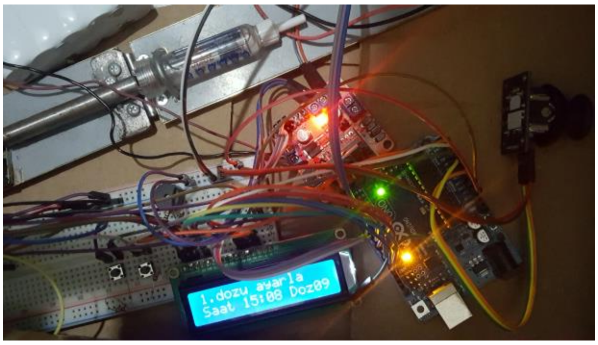


Figure 5. Bilensulin display during dose adjustment

When the insulin cartridge is finished, there is also a mechanism to switch the device to its initial position and to change the cartridge. The reservoir in which the cartridge is placed is large enough for all insulin cartridges. All possible types of insulin cartridges that are planned to be used for possible changes may be placed in this reservoir. This is a technical element that relaxes the patient both economically and in terms of ease of use. The clock module, which is located in the system, is designed in order to prevent any mistakes occurred in local clock changes or intercontinental travels, so that the patient can change it by his/her own skill or under the supervision of an individual.

3. Method of Dose Accuracy Testing

The dose of insulin injected by BS was tested by comparing with the brand insulin pens used in the marketing. The amounts of insulin injected was measured by using a sensitive balance in according to the international standard [15, 16]. Dosing accuracy was tested at 9U, 18U, 27U, 36U and 45U doses for the same type of insulin. Three insulin pens were used for comparation. The pens used in this study were the newest marketed model/technology of their brands. At the each dose, the pen was tested five times. The measurements were made by using a sensitive balance (Sartorius, ME235P-SD) calibrated before. Its measurement accuracy is 0,01mmg (the accreditation code of the calibration firm: AB-0039-K, the calibration certificate number: M15060806).

A new needle was mounted on each insulin pen according to the manufacturer's instructions before each dose. During injection, the discharging time was determined from manufacturer user manuals. The discharging of insulin must be taken for 5-10 seconds. Insulin can still be flowing out of the pen for several seconds after the button is fully depressed because of the mechanics of pen devices. Weigh measurements were taken immediately after each dose discharge in order to reduce the potential of evaporation of the expelled insulin. Environmental conditions were kept constant.

For 9U testing, the measurement was repeated five times for each insulin pen, resulting in a total of 15 measurements for three insulin pen. The same procedure was repeated for 18U, 27U, 36U and 45U. Totaly, 75 measurements were obtained. In order to eliminate potential user affect, all doses were delivered by the same person.

The related ISO standard is ISO 11608-1:2000. Specified limits based on ISO Standard were ±1 unit for a 10 unit dose and ±1,5 units for a 30 unit dose [16].

4. Results

The insulin amounts pumped in each round of the motor of BS are given in Table 1. Each measurement was repeated five times with tare measurements taken into account. The average of these five measurements were determined as the amount of each turn pumping. Again, the standard deviations of measurements were calculated.

Table 1. Amounts of insulin given in each round of the BS motor.

T	Tare (g)	Total Weight (g)	Net Weight (g)	Mean (g)	std
1	1,09899	1,17950	0,08051	0,088574	0,00526
	1,11212	1,20273	0,09061		
	1,10954	1,20170	0,09216		
	1,10886	1,20224	0,09338		
	1,10268	1,18889	0,08621		
2	1,10838	1,29720	0,18882	0,187832	0,00570
	1,10342	1,29965	0,19623		
	1,09886	1,28767	0,18881		
	1,11046	1,29446	0,18400		
	1,10652	1,28782	0,1813		

T	Tare (g)	Total Weight (g)	Net Weight (g)	Mean (g)	std
3	1,11169	1,39953	0,28784	0,29081	0,00767
	1,11308	1,41704	0,30396		
	1,10808	1,39294	0,28486		
	1,11128	1,40211	0,29083		
	1,11986	1,40642	0,28656		
4	1,10348	1,48964	0,38616	0,386978	0,00194
	1,10744	1,49558	0,38814		
	1,11320	1,49824	0,38504		
	1,11244	1,50221	0,38977		
	1,10989	1,49567	0,38578		
5	1,11317	1,59411	0,48094	0,485158	0,00708
	1,10325	1,59140	0,48815		
	1,11300	1,59438	0,48138		
	1,11011	1,58920	0,47909		
	1,10655	1,60278	0,49623		

The insulin amounts injected from the most commonly used three types of brand insulin pencils in the market can be seen in Table 2. The objective is to determine the relationship between the insulin doses injected by BS and other insulin pens.

When Table 3 are examined, it can be seen that the amount of insulin pumped in each motor tour in the BS corresponds to 9U dose in the insulin pens.

Table 3. Comparison of dosages of BS and three different insulin pens.

	BS (g)	1st pen (g)	2nd pen (g)	3rd pen (g)
9U	0,088574	0,089138	0,098554	0,095342
18U	0,187832	0,188493	0,198200	0,194522
27U	0,29081	0,269786	0,279436	0,275446
36U	0,386978	0,361322	0,369500	0,365570
45U	0,485158	0,448696	0,459684	0,455278

In Figure 5, the doses injected from BS and other insulin pens were compared by using graphic method. Figure 5 shows that BS injected insulin amounts close to the doses of other insulin pens, at all doses.

Table 2. Mean delivered doses of 9U, 18U, 27U, 36U and 45U insulin with three different insulin pens. D:dose

Dose	1st pen (g)	2nd pen (g)	3rd pen (g)
9U	0,08808	0,09782	0,09472
	0,08952	0,09964	0,09518
	0,08994	0,09655	0,09492
	0,08886	0,09984	0,09602
	0,08929	0,09892	0,09587
Mean	0,089138	0,098554	0,095342
std	0,000709	0,001371	0,000577
18U	0,189820	0,19767	0,19442
	0,187933	0,19894	0,19390
	0,188257	0,19698	0,19488
	0,189255	0,19922	0,19436
	0,187198	0,19819	0,19505
Mean	0,188493	0,19820	0,194522
std	0,001047	0,000916	0,000456
27U	0,27034	0,28004	0,27523
	0,26843	0,27983	0,27621
	0,26913	0,27827	0,27448
	0,27121	0,28109	0,27572
	0,26982	0,27795	0,27559
Mean	0,269786	0,279436	0,275446
std	0,001073	0,001306	0,000644
36U	0,36215	0,37120	0,36586
	0,36098	0,36879	0,36448
	0,36244	0,37034	0,36604
	0,35986	0,36772	0,36532
	0,36118	0,36945	0,36615
Mean	0,361322	0,36950	0,36557
std	0,001026	0,001348	0,000688
45U	0,44876	0,46024	0,45572
	0,44871	0,45986	0,45328
	0,44856	0,45879	0,45604
	0,45124	0,46008	0,45628
	0,44621	0,45945	0,45507
Mean	0,448696	0,459684	0,455278
std	0,00178	0,000581	0,001206

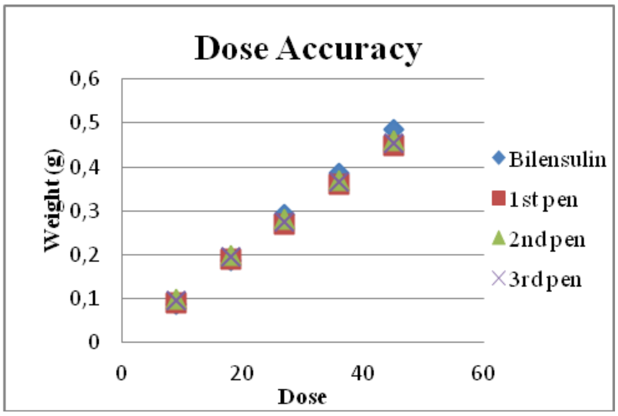


Figure 5. Comparison of mean doses for BS and the other three brand insulin pens

5. Conclusions

In this study, an easy-to-use electronic insulin pen, named as Bilensulin, was designed to prevent the wrong dose for patients with mental or visual disturbances or hand immobilities, considering the physical and memory features of such insulin systems. The dose measurements performed after design has shown that the dose accuracy of the amount of insulin injected by the BS is acceptable.

BS is an electronic insulin pen that focuses on zeroing off the error with a system that evokes the smart technology by abstracting the patient's dose completely from its own will. In addition to having a memory record like others, it also tends to keep the patient under control and reduce the risk to a minimum with an audible warning system.

The developed cartridge exchange mechanism offers an advantage in patients who have difficulty using his/her hand in terms of ease of use. In addition, the thick structure of BS allows to be easily held by such patients. The BS is different in this design structure from other insulin pens patented.

The patent of EP1095668B1 is related to the electronic insulin pen which can record the information such as insulin dosage, date and time [17]. In the patent documents numarated as US8298194 and US5688251, easy and safe methods for refilling cartridges of manually adjustable insulin pens have been explained [18, 19]. BS includes both features. It serves easy and safe methods for refilling cartridges in electronic insulin pens which adjusts dosage and time.

For this designed electronic insulin pen, the patent has been applied by putting the cartridge change mechanism into the foreground. Work is continuing to develop the device.

Acknowledgments

The patent application has been supported by The Research Fund of the University of Istanbul. The project number is FMP-2016-22534 (Property Patent Project).

6. References

[1] P. Sonksen, J. Sonksen, "Insulin: understanding its action in health and disease", *Br J Anaesth*, vol. *85*, no. 1, pp. *69-79*, July, 2000.
[2] M. P. Baruah, "Insulin pens: the modern delivery devices", *J Assoc Physicians India*, vol. *59*, pp. *38-40*, Apr, 2011.
[3] D. K. Dang, J. Lee, "Analysis of symposium articles on insulin pen devices and alternative insulin delivery methods", *J Diabetes Sci Technol.*, vol. *4*, no. *3*, pp. *558-561*, May, 2010.
[4] S. L. Schwartz, D. A. Ignaut, J. N. Bodie, "Humalog KwikPen: an insulin-injecting pen designed for ease of use", *Expert Rev Med Devices*, vol. *7*, no. *6*, pp. *735-743*, Nov, 2010.
[5] E. K. McCoy, B. M. Wright, "A review of insulin pen devices", *Postgraduate Medicine*, vol. *122*, no. *3*, pp. *81-88*, May, 2010.
[6] B. M. Rollins, D. C. Lee, M. A. Silva, "Insulin dosage preparation and administration: Prefilled pens versus syringes", *J Pharmacy Practice and Research*, vol. *43*, no. *3*, pp. *198-201*, Sept, 2013.
[7] A. Pfützner, T. Bailey, C. Campos, D. Kahn, E. Ambers, M. Niemeyer, G. Guerrero, D. Klonoff, I. Nayberg, "Accuracy and preference assessment of prefilled insulin pen versus vial and syringe with diabetes patients, caregivers and healthcare professionals", *Curr Med Res Opin*, vol. *29*, no. *5*, May, 2013.
[8] G. Spollett, "Insulin devices: addressing barriers to insulin therapy with the ideal pen", *Diabetes Educator*, vol. *34*, no. *6*, pp. *957-960, 963,967*, Nov, 2008.
[9] D. A. Ignaut, M. R. Opincar, P. E. Clark, M. K. Palaisa, S. M. Lenox, "Engineering study comparing injection force and dose accuracy between two prefilled insulin injection pens", *Curr Med Res Opin*, vol. *25*, no. *12*, pp. *2829-2833*, Dec, 2009.
[10] T. L. Pearson, "Practical aspects of insulin pen devices", *J Diabetes Sci Technol.*, vol. *4*, no. *3*, pp. *522-531*, May, 2010.
[11] A. Bode, "Development of the SoloSTAR insulin pen device: design verification and validation", *Expert Opin Drug Del*, vol. *6*, no. *1*, pp. *103-112*, Dec, 2009.
[12] D. A. Nadeau, C. Campos, M. Niemeyer, T. Bailey, "Healthcare professional and patient assessment of a new prefilled insulin pen versus two widely available prefilled insulin pens for ease of use, teaching and learning", *Curr Med Res Opin*, vol. *28*, no. *1*, pp. *3-13*, Jan, 2012.
[13] D. Burton, M. Uslan, "Diabetes and visual impairment: Are insulin pens accessible?", *AFB AccessWorld Magazine*, vol. *7*, no. *4*, July, 2006.
[14] W. J. R. R. Venekamp, L. Kerr, S. A. Dowsett, P. A. Johnson, D. Wimberley, C. McKenzie, J. Malone, Z. Milicevic, "Functionality and acceptability of a new electronic insulin injection pen with a memory feature", *Curr Med Res Opin*, vol. *22*, no. *2*, pp. *315-325*, Jan, 2006.
[15] T. Asakura, H. Seino, M. Kageyama, N. Yohkoh, "Dosing accuracy of two insulin pre-filled pens", *Curr Med Res Opin*, vol. *24*, no. *5*, pp. *1429-1434*, Apr, 2008.
[16] *Pen-injectors for Medical Use - Part 1: Pen-injectors - Requirements and Test Methods,* ISO 11608-1:2000, 2000.
[17] *Electronic medication delivery pen having a multifunction actuator,* Patent number: EP1095668B1, 2007.
[18] *Injection device and a method of changing a cartridge in the device,* Patent number: US 8298194, 2012.
[19] *Cartridge loading and priming mechanism for a pen injector,* Patent number: US5688251, 1997.

Mana Sezdi (Bandirma, Turkey, 1972) graduated from Uludag University as an electronical engineer in 1994. She received her master and Ph.D. on Biomedical Engineering from Bogazici University, in 1998 and 2005, respectively. Between 1996-2006, she worked as researcher in Biomedical Engineering Institute, Bogazici University in Turkey. She has been working in the Biomedical Device Technology Program of Istanbul University as instructor since 2007. She is also assistant director in Biomedical and Clinical Engineering Department of the Istanbul University's Medicine Faculties; Cerrahpasa Faculty of Medicine and Istanbul Faculty of Medicine since 2007. She is an author of more than 85 publications in total, including 1 book, 4 book chapters and 80 research papers in international refereed journals, national and international conference proceedings. Her book is "The quality of stored human blood" that was published by Lambert Academic Publishing in 2011. Her published book chapters; "Dose Optimization for the Quality Control Tests of X-Ray Equipment" in the book of "Modern Approaches to Quality Control" (Croatia, Intech, 2011) and "Medical Technology Management and Patient Safety" in the book of "A Broadmap of Biomedical Engineers and Milestones" (Croatia, Intech, 2012). Her research interests include biomedical instrumentation and measurements, clinical engineering, medical calibration measurements and performance tests of diagnostic imaging systems. She is the member of Association of Biomedical and Clinical Engineering in Turkey. Between 1996 and 2006, she was in the program committee of National Biomedical Conference (BIYOMUT). Since 2010, she has been serving on program committee of Health Technologies National Conference (TIPTEKNO Conferences).

Veli BAYSAL et al. / IU-JEEE Vol. 17(1), (2017), 3081-3085

BLOCKING OF WEAK SIGNAL PROPAGATION VIA AUTAPTIC TRANSMISSION IN SCALE-FREE NETWORKS

Veli BAYSAL [1], Ergin YILMAZ [1], Mahmut ÖZER [2]

[1]Department of Biomedical Engineering, Engineering Faculty, Bülent Ecevit University, 67100 Zonguldak, Turkey
[2]Department of Electrical and Electronics Engineering, Engineering Faculty, Bülent Ecevit University, 67100 Zonguldak, Turkey
veli.baysal@beun.edu.tr, erginyilmaz@yahoo.com, mahmutozer2002@yahoo.com

Abstract: *In this paper, the effects of autapse, a kind of synapse formed between the axon or soma of a neuron and its own dendrite, on the transmission of weak signal are investigated in scale-free neuronal networks. In the study, we consider that each neuron has an autapse modelled as chemical synapse. Then, a weak signal that is thought to carry information or an unwanted activity such as virus is applied to all neurons in the network. It is seen that the autapse with its small conductance values can slightly increases the transmission of weak signal across the network when the autaptic time delay is equal to the intrinsic oscillation period of the Hodgkin-Huxley neuron. Interestingly, when the autaptic time delay becomes equal to half of this intrinsic period or its integer multiples the autapse can prominently blocks the weak signal transmission. Also, as the autaptic conductance is increased the weak signal transmission is completely impeded by the autapse with its proper auatptic time delays. One consider that the weak signal is an unwanted or virius threatening the whole network, this autaptic mechanism is an efficient way to protect the network from attacks.*
Keywords: *Autapse, Scale-free, Blocking of weak signal.*

1. Introduction

Information exchange among neurons is fulfilled via special structures called synapse. There are two different types of synapses: electrical synapses and chemical synapses [1]. Synaptic connections are commonly occurred between two different neurons. On the other hand, a different type of synaptic connection called autapse, which is established between the axon and the dendrites of the same neuron, was firstly introduced by Vander Loos and Glaser [2]. Autapse could be electrical synapse or chemical synapse [3, 4]. Presence of this synaptic connection in different brain regions was uncovered various experimental studies by using different experimental techniques [5-11]. Tamas et al. showed that neurons in visual cortex could have roughly between 10 to 30 inhibitory autapses [11]. Lübke at al demonstrated that the 80 percent of cortical pyramidal neurons have autaptic connections in neocortex of human brain [5]. Bacci et al reported that GABAergic autaptic activity is present in fast-spiking interneurons of layer V in neocortical slices. Also they demonstrated that autaptic activity has significant inhibitory effect on the repetitive firing, and can increase current threshold for evoking action potential [12].

In addition to above studies where the presence of autapse have been shown with experimental studies, there are some studies investigating the effects of autapse on neuronal dynamics[13-24]. Saada et al. showed that autapse can cause persistent activity in B31/B32 neurons of Aplysia [13]. Bacci and Huguenard indicated that autapse can have determinative effect on the spike time of interneurons in neocortical slices [14]. Li et al. showed via histogram analysis that the number of spikes in stochastic Hodgkin-Huxley neuron is decreased in the presence of autapse [15]. Autapse can trigger the formation of spiral wave in regular network comprised of Hindmarsh-Rose (HR)[16]. Masoller et al. [17] studied how the subthreshold dynamics of Hodgkin-Huxley (HH) neuron interacts with time-delayed feedback and noise. They reported that for negative feedback, the firing rate can be lower than in the noise-free situation, for positive feedback, there are regions of delay values where the noise-induced spikes are inhibited by the feedback (i.e., autapse). Connelly found that autapse enhances the synchrony of basket cell membrane potentials across the network during neocortical gamma oscillations [18]. Wang et al studied that autapse-induced transition of firing pattern using the HR neuron model theoretically. They indicated that delayed autaptic feedback connection switches the electrical activitie of the HR neuron among quiescent, periodic and chaotic firing patterns [19]. In Ref [20], it was shown that the autapse can enhance or abolish the status of mode-locking and can effectively regulate the neuronal response. Sainz-Trapaga et al. investigated the dynamics of thermally sensitive neurons that display intrinsic oscillatory activity. They reported that a self-feedback causes spikes by

Received on: 31.01.2017
Accepted on: 14.03.2017

increasing the amplitude of the subthreshold oscillations above the threshold [21]. In Ref [22], it was shown that single spikes and burst type spikes is a sensitive function of autaptic time delay. Besides, Yılmaz et al. revealed that the presence of autapse can significantly enhances the propagation of pacemaker activity across both scale-free (SF) and small world (SW) neuronal networks [23-24].

In the above studies conducted in network level, it is considered that only pacemaker neuron has autaptic connection. But in realistic conditions, many neurons in the network can have this type of connection. In this study, we take into account that all neurons in the network have autapse modeled as chemical synapse. A weak signal which can be thought an unwanted signals (may be virus or an anomaly) is injected to all neurons. Then, the effects of autapse on the transmission or propagation of this weak signal is investigated in scale-free neuronal networks. When the obtained results evaluated, we briefly say that, autapse can become an efficient control mechanism to prevent the spreading of unwanted signals in scale-free neuronal networks.

2. Model and Methods

In order to simulate the stochastic neuronal dynamics in the scale-free network effectively, we employ the Hodgkin-Huxley equations [25].

$$C_m \frac{dV_i}{dt} + g_K^{max} n_i^4 (V_i - E_K) + g_{Na}^{max} m_i^3 h (V_i - E_{Na}) + g_l (V_i - E_l) = I_{inj} - I_i^{aut} + \sum_{j=1}^{N} \varepsilon_{ij} (V_j - V_i),\ i = 1, 2, \dots N \tag{1}$$

where $C_m = 1\mu F/cm^2$ is the capacity of the cell membrane, V_i denotes the membrane potential of neuron i. $g_{Na}^{max} = 120 mScm^{-2}$ and $g_K^{max} = 36 mScm^{-2}$ respectively denote the maximal potassium and sodium conductance, when all ion channels are open. The leakage conductance is assumed to be constant, equaling $g_l = 0.3$. $E_K = -77$ mV, $E_{Na} = 50$ mV and $E_l = -54.4$ mV are the reversal potentials for the potassium, sodium and leakage current, respectively. N is total number of neuron in the netwoks. In this paper it is assumed that $\varepsilon_{ij} = \varepsilon$, if the neurons *i* and *j* are connected; otherwise $\varepsilon_{ij} = 0$. Here, I_{inj} is given with the following equation [15]:

$$I_{inj} = \sin(0.3t) \tag{2}$$

I_i^{aut} is the autaptic current stemming from the autaptic connection of neuron i. Autapse is assumed as chemical synapse in this paper and modeled using the so-called fast threshold modulation given by the following function.

$$I_i^{aut} = -\kappa (V_i(t) - V_{syn}) S(t - \tau) \tag{3}$$

$$S(t - \tau) = 1/\{1 + \exp(-k(V_i(t - \tau) - \theta))\} \tag{4}$$

where κ denotes the conductance of line that is flowed autaptic current on, and τ represents the autaptic time delay, which occurs because of the finite propagation speed during axonal transmission. $V_{syn} = 2$ mV for excitatory chemical autapse, $k = 8$ and $\theta = -0.25$.

m_i and h_i represent the activation and inactivation variables for sodium channels of neuron i, respectively. The activation variables for potassium channels of neuron of i is expressed with n_i. The gating dynamics is described by the Langevin generalization that based on Fox's algorithm as follows [26]:

$$\frac{dx}{dt} = \alpha_x(V)(1 - x) - \beta_x(V) x + \zeta_x(t),\ x = m, n, h \tag{5}$$

where $\alpha_x(V)$ and $\beta_x(V)$ are the voltage-dependent rate functions for the gating parameter x [25].

$$\alpha_m(V) = \frac{0.1(V+40)}{1-\exp(-(V+40)/10)} \tag{6}$$

$$\beta_m(V) = 4\exp[-(V + 65)/18] \tag{7}$$

$$\alpha_h(V) = 0.07\exp[-(V + 65)/20] \tag{8}$$

$$\beta_h(V) = \frac{1}{1+\exp[-(V+35)/10]} \tag{9}$$

$$\alpha_n(V) = \frac{0.01(V+55)}{1-\exp[-(V+55)/10]} \tag{10}$$

$$\beta_h(V) = 0.125\exp[-(V + 65)/80] \tag{11}$$

ζ_x denotes the independent zero mean Gaussian white noise whose autocorrelation functions are given as follows [25]:

$$\langle \zeta_m(t) \zeta_m(t') \rangle = \frac{2\alpha_m \beta_m}{N_{Na}(\alpha_m + \beta_m)} \delta(t - t') \tag{12}$$

$$\langle \zeta_h(t) \zeta_h(t') \rangle = \frac{2\alpha_h \beta_h}{N_{Na}(\alpha_h + \beta_h)} \delta(t - t') \tag{13}$$

$$\langle \zeta_n(t) \zeta_n(t') \rangle = \frac{2\alpha_n \beta_n}{N_K(\alpha_n + \beta_n)} \delta(t - t') \tag{14}$$

where N_{Na} and N_K represent the total numbers of sodium and potassium channel, and calculated as $N_{Na} = \rho_{Na} S$ and $N_K = \rho_K S$, respectively. S is the cell size or the membrane area used for the scaling of channel noise intensity. The number of channels per square micrometer of cell size is $\rho_{Na} = 60\ \mu m^{-2}$ for sodium and $\rho_K = 18\ \mu m^{-2}$ for potassium. It is given in Eq. (12, 13, 14) that when the cell size is large enough the stochastic effect added by the ion channels to the membrane potential is trivial, but when the cell size is small the stochastic effect due to the ion channels is very crucial [31].

Following the procedure in [28], we construct the scale-free neuronal network, using N=200 neurons with different average degree of connectivity, k_{avg}. To quantitatively demonstrate the weak signal propagation degree, we calculate Fourier series coefficients. To do so, we first calculate the average membrane potential $V_{avg}(t) =$

$\frac{1}{N}\sum_i^N V_i(t)$ during N = 1000 periods. Then, we calculate the Fourier coefficients as follows:

$$Q_{sin} = \frac{\omega}{2N\pi}\int_0^{2N\pi/\omega} 2V_{avg}(t)(t)\sin(\omega t)\, dt \quad (15)$$

$$Q_{cos} = \frac{\omega}{2N\pi}\int_0^{2N\pi/\omega} 2V_{avg}(t)\cos(\omega t)\, dt \quad (16)$$

$$Q = \sqrt{Q_{sin}^2 + Q_{cos}^2} \quad (17)$$

where, $\omega = 2\pi/t_s$ is the frequency of the weak signal. Notably, the larger the Q the better the weak signal propagation.

3. Results and Discussion

In all previous studies where the propagation of the weak localized pacemaker activity is considered, only one neuron acting as pacemaker has an autapse. But, here we consider that each neuron in the network has one autapse modeled as chemical synapse. Then, we investigate the effects of autapse on the transmission or propagation of weak signal applied to the all neurons. To do so, we initially fix the cell size $S = 16\mu m^2$ and the average degree of connectivity $k_{avg} = 10$ and the coupling constant $\varepsilon = 0.05$. In Fig.1, we give the dependence of Q on the autaptic time delay for low levels of autaptic conductances. Also to make a comparison, we demonstrate Q values of the network in the absence of autapse (black straight line in Fig.1).

It is seen in Fig.1 that the weak signal propagation capacity of the network slightly increases for finely tuned τ. But, interestingly, when τ is equal to half of the intrinsic oscillation period of HH neuron ($T_{osc} \approx 21ms$ [23]) or its odd multiples, the weak signal propagation throughout network decreases prominently compared with the without autapse.

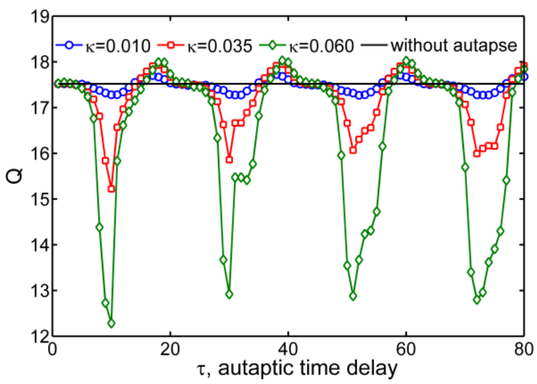


Figure 1. Effect of low autaptic conductance levels on the transmission of weak signal (ε=0.05, S=16 μm², N=200, k_{avg}=10)

To provide clear evidence for the results in Fig.1, we give the average membrane potential and the weak signal in the same panel for three different autaptic time delay values when autaptic conductance $\kappa = 0.06$ in Fig 2.

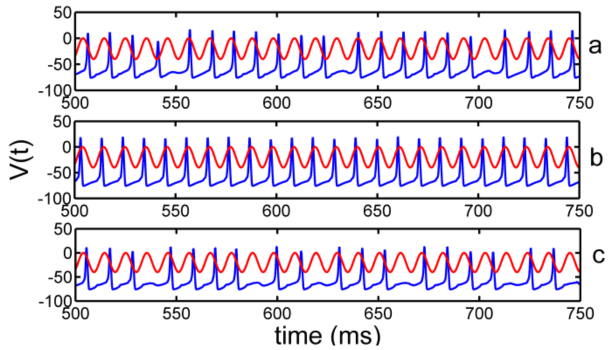


Figure 2. Average membrane potential at different autaptic time delay value with weak signal (weak signal is magnified 20 times and is shifted towards below20 units at vertical axis) a) τ = 30 ms b) τ = 17 ms c) τ = 10 ms (ε=0.05, S=16 μm², N=200, k_{avg}=10, κ = 0.06)

It is seen that the overlap between the weak signal and average membrane potential is maximum, and the average membrane potential fires when the weak signal is maximum. But in Fig 2a and Fig. 2c, matching between weak signal and average membrane potential is disrupted. Particularly in Fig 2c, the average membrane potential spikes do not occur at the time when the weak signal takes the peak value, and cycle skipping occurs. As a consequences, when the match between the weak signal and average membrane potential is well, the obtained Q values are high, which indicates better propagation of weak signal across the network. If the match between average membrane potential and the weak signal is bad and cycle skipping occurs, low Q values are obtained.

In Fig. 3, we show the dependence of Q on τ values for intermediate and high level of autaptic conductance levels. As seen in Fig.3, as the autaptic conductance level increases the propagation of weak signal across the network reduces, and even, at a strong autaptic conductance level ($\kappa = 0.76$) the propagation of weak signal is ceased by autapse for some autaptic time delay intervals when compared to the without autapse. Interestingly, when the auatptic time delay equals to the intrinsic oscillation period (T_{osc}) of HH neuron or its integer multiples the level of weak signal propagation takes the values roughly equal to the ones obtained in the absence of autapse.

To provide more evidence about the blockage of weak signal transmission, we plot the average membrane potential and the weak signal in the same panel for different autaptic time delay values in Fig 4.

As seen in Fig.4a, average membrane potential of the network has spikes occurring at approximately time instances when the weak signal has peak value. This coherence between average membrane potential and weak signal causes high Q values. But, in Fig.4b, the spike times of average membrane potential match the negative peak of the weak signal, that is, the synchronization between weak signal and spiking activity is destroyed, which leads to obtain small Q values.

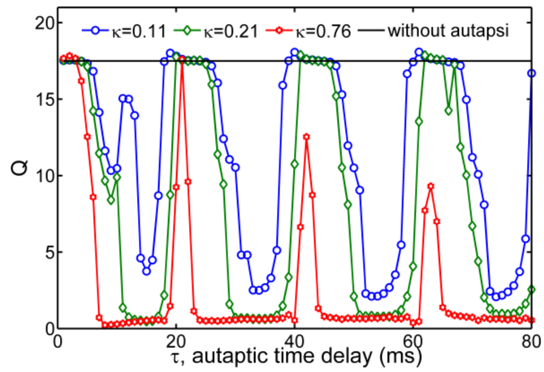


Figure 3. Effect of autaptic conductance on transmisson of weak signal depend on autaptic time delay (ε=0.05, S=16 μm^2, N=200, k_{avg}=10)

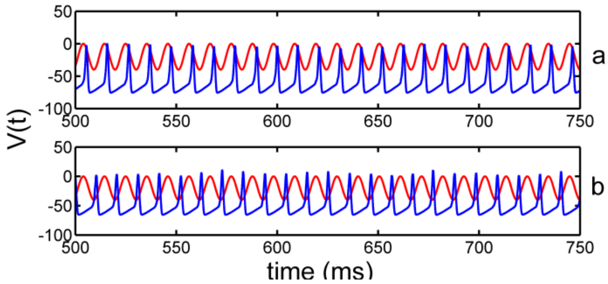


Figure 4. Average membrane potential and weak signal (weak signal is magnified 20 times and is shifted 20 units at vertical axis) for different auatptic time delays. a) $\tau = 21$ ms b) $\tau = 35$ ms (ε=0.05, S=16 μm^2, N=200, k_{avg}=10, $\kappa = 0.21$)

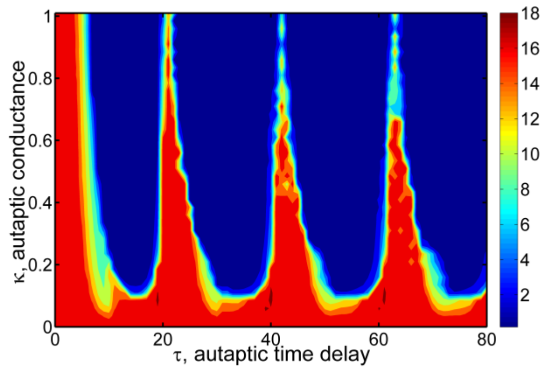


Figure 5. The dependence of Q on autaptic conductance and autaptic time delay (ε=0.05, S=16 μm^2, N=200, k_{avg}=10)

To get a global view, we show the contour plot of Q on $\kappa - \tau$ parameters space in Fig. 5. Results reveal that for low values of autaptic time delay (roughly $\tau < 10\ ms$), the weak signal propagation across the network is not affected by the variations in autaptic conductance, and the Q values take approximately the same value obtained in the absence of autapse. Similarly, when the autaptic conductance is lower than $\kappa = 0.1$, there is not any effect of autaptic time delay on the propagation of weak signal in the network. When $\tau > 10\ ms$ and $\kappa > 0.1$, we obtain different resonance islands where the degree of propagation of weak signal is almost the same with that obtained in the absence of auatpse (red shaded region). Outside of these resonance islands, we obtain that the presence of autapse significantly blocks the propagation of weak signal.

4. Conclusions

In sum, the effects of autapse on the propagation of weak signal are investigated in scale-free neuronal networks where each neuron has a chemical autapse. We obtain that when each neuron has autaptic connection in the network, the presence of autapse does not augment the propagation of weak signal in contrast it prevents the propagation of weak signal in the network. If someone assumes that this weak signal carries an unwanted signal such as infectious disease, virus, schizophrenic signal etc., the presence of autapse will be an efficient way to cope with this unwanted disturbances.

5. References

[1] EJ. Furshpan, DD. Potter, “Transmission at the giant motor synapses of the crayfish.", *J. Physiol*., vol. *145*, no. *2*, pp. *289-325*, March 1959.

[2] H. Van der Loos, E.M. Glaser, “Autapses in neocortex cerebri: synapses between a pyramidal cell’s axon and its own dendrites.", *Brain Res*., vol. *48*, no. *24*, pp. *355-360*, December 1972.

[3] JM. Bekkers, “Neurophysiology: are autapses prodigal synapses?”, *Curr Biol*., vol. *8*, no. *2*, pp. *52-55*, January 1998.

[4] K. Yamaguchi, “Autapse Encyclopedia of Neuroscience" (M. D. Binder ed.), N Hirokawa and U Windhorst (Springer Berlin Heidelberg) 2008.

[5] J. Lübke, H. Markram, M Frotscher, B. Sakmann, “Frequency and dendritic distribution of autapses established by layer 5 pyramidal neurons in the developing rat neocortex: Comparison with synaptic innervation of adjacent neurons of the same class.", *J. Neurosci*., vol. *16*, no.*10*, pp *3209-3218*, May 1996.

[6] M. H. Flight, “Neuromodulation: Exerting self-control for persistence.", *Nat. Rev. Neurosci*., vol. *10*, no. *316*, May 2009

[7] T. Branco, K. Staras “The probability of neurotransmitter release: variability and feedback control at single synapses.", *Nat. Rev. Neurosci*., vol. *10*, pp. *373-383*, May 2009.

[8] A.B. Karabelas, D.P. Purpura, “Evidence for autapses in the substantia nigra" *Brain Res*., vol. *200*, no. *2*, pp. *467-473*, November 1980.

[9] M.R. Park, J.W. Lighthall, S.T. Kitai, “Recurrent inhibition in the rat neostriatum.", *Brain Res*., vol. *194*, no *2*, pp. *359-369*, August 1980.

[10] F. Kimura, Y. Otsu, T. Tsumoto, “Presynaptically silent synapses: spontaneously active terminals without stimulus-evoked release demonstrated in cortical autapses.”, *J. Neurophysiol*., vol. *77*, no. *5*, pp. *2805-2815*, May 1997.

[11] G. Tamás, E.H. Buhl, P. Somogyi, “Massive autaptic self-innervation of GABAergic neurons in cat visual cortex", *J. Neurosci*., vol. *17*, no. *16*, pp *6352-6364*, August 1997.

[12] A. Bacci, J.R. Huguenard, D.A. “Prince Functional autaptic neurotransmission in fast-spiking interneurons: a novel form of feedback inhibition in the neocortex”, *J. Neurosci*., vol. *23*, no *3*, pp. *859–866*, February 2003.

[13] R. Saada, N. Miller, I. Hurwitz, A. J. Susswein, “Autaptic excitation elicits persistent activity and a plateau potential in a neuron of known behavioral function.", *Curr. Biol*., vol. *19*, no. *6*, pp. *479-484*, March 2009.

[14] A. Bacci, J.R. Huguenard, “Enhancement of Spike-Timing Precision by Autaptic Transmission in Neocortical Inhibitory

Interneurons", *Neuron*, vol. *49*, no. *1*, pp. *119-130*, January 2006.
[15] Y. Li, G. Schmid, P. Hänggi, L. Schimansky-Geier, "Spontaneous spiking in an autaptic Hodgkin-Huxley setup", *Phys. Rev. E.*, vol. *82*, no. *6*, pp. *061907*, December 2010.
[16] H.X Qin, J. Ma, C. Wang, R. Chu, "Autapse-induced target wave, spiral wave in regular networkof neurons", *Physics, Mechanics and Astronomy*, vol. *57*, no. *10*, pp. *1918-1926*, October 2014.
[17] C. Masoller, M.C. Torrent, J. García-Ojalvo, "Interplay of subthreshold activity, time-delayed feedback, and noise on neuronal firing patterns" *Phys. Rev. E*, vol. *78*, no. *4*, pp. *041907*, October 2008.
[18] W.M. Connelly, "Autaptic connections and synaptic depression constrain and promote gamma oscillations" *PLoS ONE*, vol. *9*, no. *2*, pp. *e89995*, February 2014.
[19] H. Wang, J. Ma, Y. Chen, Y. Chen, "Effect of an autapse on the firing pattern transition in a bursting neuron." *Commun. Nonlinear Sci. Numer. Simul.*, vol. *19*, no. *9*, pp. *3242–3254*, September 2014.
[20] H. Wang, Y. Sun, Y. Li, Y. Chen, "Influence of autaptic self-feedback on mode-locking structure of a Hodgkin–Huxley neuron under sinusoidal stimulus.", *J. Theoret. Biol.*, vol. *358*, pp. *25–30*, May 2014.
[21] M. Sainz-Trapaga, C. Maseller, H.A. Braun, M.T. Huber "Influence of time-delayed feedback in the firing pattern of thermally sensitive neurons.", *Phys. Rev. E*, vol. *70*, no. *3*, pp. *031904*, September 2004.
[22] GC Sethia, J Kurths, A Sen, "Coherence resonance in an excitable system with time delay.", *Physics Letters A*, vol. *364*, no. *3*, pp. *227-230*, April 2007.
[23] E. Yılmaz, V. Baysal, M. Ozer, M. Perc, "Autaptic pacemaker mediated propagation of weak rhythmic activity across small-world neuronal networks.", *Physica A*, vol. *444*, pp. *538-546*, February 2016.
[24] E. Yılmaz, V. Baysal, M. Perc, M. Ozer, "Enhancement of pacemaker induced stochastic resonance by an autapse in a scale-free neuronal network.", *Science China Technological Sciences*, vol. *59*, no.*1*, pp. *1–7*, March 2016.
[25] E. Yılmaz, M. Ozer, V. Baysal, M. Perc, "Autapse-induced multiple coherence resonance in single neurons and neuronal networks.", *Scientific Reports*, vol. *6*, pp. *30914*, August 2016.
[26] A L. Hodgkin, A F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve.", *J Physiol*, vol. *117*, no. *4*, pp. *500–544*, August 1952.
[27] R.F. Fox, "Stochastic versions of the Hodgkin–Huxley equations.", *Biophys. J.*, vol. *72*, pp. *2068–2074*, no. *5*, May 1997.
[28] A L. Barabasi, R. Albert "Emergence of Scaling in Random Net-works." *Science*, 1999, 286: 509-512.

INVESTIGATION OF DNA-MAGNETIC NANOPARTICLE INTERACTION BY MAGNETO-OPTICAL LINEAR DICHROISM

Selma Önal KÖKLÜ[1], İsmail YARİÇİ [1*], Yavuz ÖZTÜRK [1]

[1]Ege University, Electrical and Electronics Engineering, İzmir, Turkey
selmaonalkoklu@gmail.com,yariciismail@gmail.com, yavuz.ozturk@ege.edu.tr

Abstract: *In this study, interaction of produced magnetic nanoparticles (MNPs) with DNA was investigated by using designed magneto-optical linear dichroism (MOLD) measurement system.* Fe_3O_4 *nanoparticles (NPs), produced by using co-precipitation method were characterized by means of structure, magnetic, and size distribution. These NPs coated with oleic acid and Tetramethylammoniumhydroxide (TMA) in water and ethanol. Change of the linear dichroism property of dissolved NPs in the fluid under applied magnetic field was observed due to the interaction of produced MNPs with DNA. Obtained results were shown that it is possible to develop a DNA sensor by using MNPs and MOLD effect.*

Keywords: *Magneto-optical linear dichroism, DNA, Magnetic nanoparticles MNPs.*

1. Introduction

Nanoparticles (NPs), defined as particles with dimensions of 100 nm and below, form the basis of nanotechnology. MNPs are promising for many areas such as optics, electronics, medicine, drug delivery, diagnosis and treatment of diseases due to the behavior against the magnetic field and their adjustable dimensions. Nanotechnology has provided significant advantages for scientists in the field of many healthcare and life sciences, both in the cellular dimension and molecular level [1-5]. Furthermore, they are biocompatible with molecules such as DNA and Protein; they can be linked to this kind of structures [6]. They also have many potential applications like cancer diagnosis and treatment, hyperthermia, DNA detection, DNA isolation, magnetic resonance imaging and tissue therapy [7].

Nowadays, biocompatible materials comprised of the combination of DNA and MNPs attract particular attention with regard to their design. The administration of these systems is extremely important for the development of areas such as nano-electronic, biomedical diagnosis and treatment (development of sensitive biosensors and effective pharmaceuticals). The combination of MNPs with biological molecules and especially nucleic acids allows the development of various nanobiohybrid systems which have unique magnetic properties and biological selectivity [8,9].

DNA-MNP interaction may show different magneto optic properties in the liquid under applied magnetic field [10]. Although the magnetic fluids (suspended MNPs in a carrier liquid) are isotropic; with the applied external magnetic field they show anisotropic characteristic and exhibit birefringence effects and linear and circular dichroism effects. Such interactions under magnetic field are called magneto-optic effects [11]. The MOLD effect is explained by the formation of chain structures of MNPs in liquid under applied magnetic field [12,13]. Every particle in the electric field under this cluster is considered to be oscillating dipoles. These dipoles interact to each other because of their closeness. This interaction is asymmetrical depending on the direction of the light beam. This asymmetry creates the optical anisotropy effect that lead to linear dicroishm and linear birefringence effect in the magnetic fluid [11].

There are a few studies focused on DNA-magnetic fluid interaction and magneto-optical effects [10]. Circular-dichroism effect is generally used to investigate molecules such as DNA and proteins [14]. However, this DNA-related measurement method only works in the ultraviolet (UV) region [14]. So it is important to develop measurement methods in the visible region of light. With this motivation, in this study, MOLD effect of the influence of single stranded and double stranded DNA-MNPs solutions was studied. For this purpose, a MOLD measurement system was designed. Afterwards, the size, structure and magnetic properties of produced MNPs were characterized. The MNPs were coated with oleic acid and TMA in order to prevent agglomeration. The coated NPs were suspended in water and then the MOLD properties of interaction with single stranded, double stranded DNA were examined.

Received on: 31.01.2017
Accepted on: 14.03.2017

2. Material and Method

2.1. Production of MNPs

The co-precipitation method was used for synthesis of MNPs. During the production, as the first step FeCl3.6H2O (0.01 mol) and FeCl2.4H2O (0.02 mol) were dissolved in pure water and in a 1M HCl prepared solution in separate cups as separate solutions. Then they are mixed with molar ratios of 1: 2. The prepared solutions were stirred in a nitrogen gas atmosphere for 30 minutes, after that in order to adjust the pH value of solution and to ensure the continuity of the synthesis; ammonium was added into the mixture. Ammonium was added until the pH value was 10, which is accepted optimum value for the solution [15]. 15 minutes after the chemical reaction had started; the solution was washed 3 times with 5% ammonium solution to prevent the surface charge density of the particles from decreasing and the particles to avoid agglomeration. Some of the magnetic particles precipitated on the bottom were suspended in pure water by coating with tetramethylammoniumhydroxide (TMA) and oleic acid.

2.2. Analysis of MNPs

Structural analysis of the synthesized NPs, size distribution of the NPs in the liquid, and magnetic and optical properties of the NPs were characterized using X-ray diffraction (XRD), dynamic light scattering (DLS) and Vibration Sample Magnetometer (VSM) devices, respectively.

2.3. MOLD Measurement System

A magneto-optic experimental setup designed for this study was used to investigate the polarization-dependent transmission changes of the magnetic liquid at different magnetic field values. Magnetic liquid were prepared to include single stranded DNA and produced magnetic liquid.

A fiber coupled tungsten lamp used to obtain wide wavelength range light source. The output of light source polarized using Glan-Thomson type polarizer. Light was passed through the lens system to the sample and spectrometer. The magnetic field was generated by an electromagnet controlled with the adjustable current source.

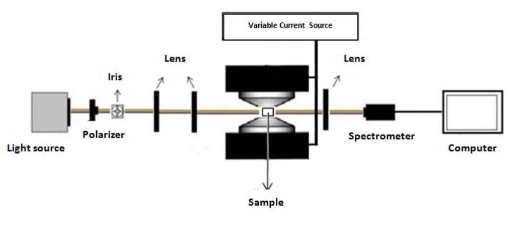


Figure 1. The designed magneto-optical measurement system

Depending on the polarization direction and the magnetic field values, the amount of transmitted light through the sample was measured at a wavelength range of 500 nm-1000 nm by USB-4000-UV-VIS spectrometer.

3. Experimantal Results

Structural analyzes of the produced MNPs were taken by a Philips Expert 1830 model XRD. The XRD result of the sample is given in Fig.2. As seen in Fig. 2, the structure of sample is Fe3O4.

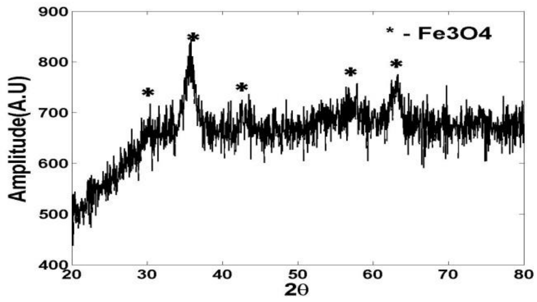


Figure 2. X- ray diffraction pattern of MNPs

The size distributions of MNPs in solution were determined using a Zetasizer 4 Nano S-Malvern DLS (Dynamic Light Scattering). According to the DLS results, size of the produced magnetic particle and the standard deviation was determined as 34 nm and 17 nm respectively.

The magnetic analysis of the NPs was performed by using Lakeshore 736, 7400 VSM instrument. As a result of the measurements, it is measured that the coercivity of NPs is low and the particles show superparamagnetic behavior. The saturation magnetization value of the particle in our study was found to be 62 emu / g. Saturation magnetization value of macro size solid Fe3O4 is 93 emu /g [16]. For NPs, the saturation magnetization is generally smaller than the solid one. One of the reasons for this is explained by the negative effect of magnetization of the non-magnetic layer on the surfaces of NPs [16].

The optical transmission test for the magnetic fluid mixture prepared with oleic acid and ethyl alcohol was carried out as shown in Fig.3. It is seen that the light transmission perpendicularly to the direction of the applied magnetic field (P90) and parallel to the magnetic field (P0) under the magnetic field is different. Thus, the MOLD effect appears clearly. In many studies on this subject, this effect is explained by the chain structure of the particles under the magnetic field [11]. Electric dipole interactions between the MNPs forming the chain structure cause different absorption.

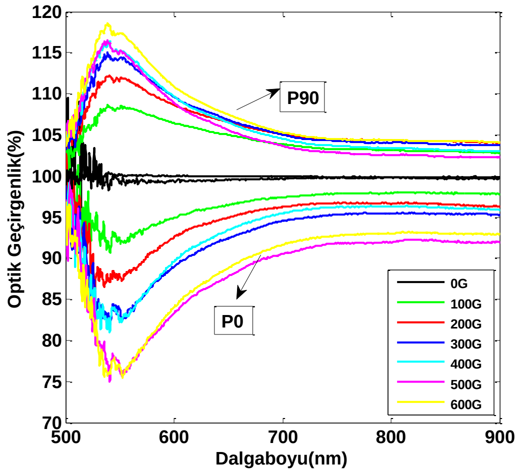


Figure 3. The optical transmission test for the magnetic fluid mixture prepared with oleic acid and ethyl alcohol was carried out

20 mg single-stranded herringbone DNA was mixed into the magnetic fluid. The optical transmittance of this mixture changed due to polarization under applied field. This optical transmittance change for different polarization under applied magnetic field was shown in Fig.4 and Fig.5. Optical transmission increases at different amounts depending on polarization under the same external magnetic fields. For example, the P0 polarization under the 400G external field has a maximum optical transmittance of 123%, but this value remains at 111% for P90. The results clearly state that the MOLD effect is different for magnetic fluid with and without single stranded DNA. The optical transmission increases for both polarizations under the applied magnetic field. The optical absorption increase which is explained by the formation of the chain structures shown in Fig.3 was not observed for the polarization of P0 in Fig.5.

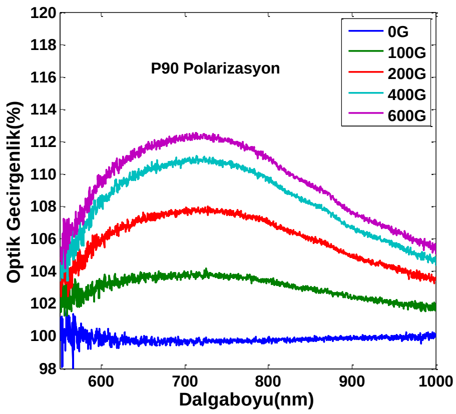


Figure 4. The wavelength-dependent optical transmission of magnetic fluid containing 20 mg single stranded DNA in P90 polarity out

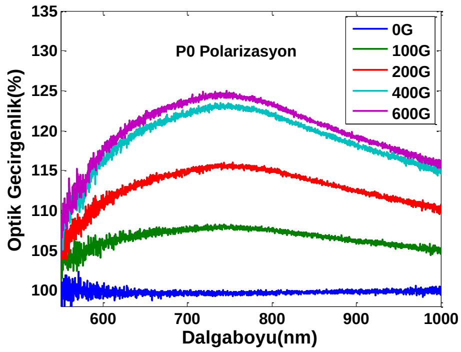
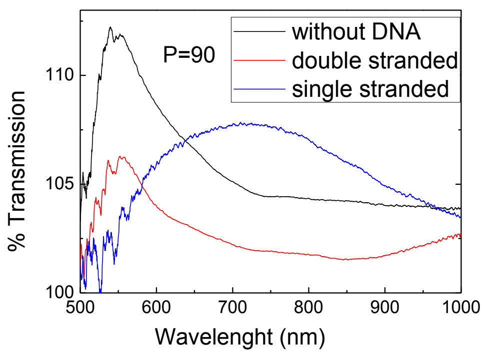


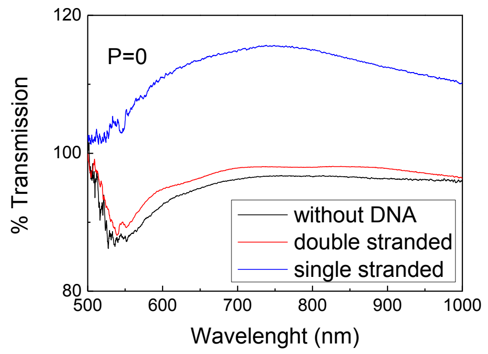
Figure 4. The wavelength-dependent optical transmission of magnetic fluid containing 20 mg single stranded DNA in P0 polarity out

The same study also was repeated for double stranded DNA and obtained results were shown in another study [15].

In order to compare the change in the transmission of light of single stranded samples, double stranded samples and samples which do not contain DNA molecules as a result of polarization, magnetic fluids which contain 20 mg single - 20 mg double stranded DNA in each and fluid samples which do not contain DNA were shown on the same graphic under 100 Gauss fixed magnetic field. Magnetic particles bind with the double stranded DNA less than the single stranded because it has a helix structure. Changes related to polarization in double stranded DNA-magnetic fluid mixture show that the chain structure is formed in the mixture. Transmission of light related to polarization is lower compared to the fluid which does not contain DNA. Changes related to polarization are considerably low in the single stranded version. This situation shows that the transmission of light related to polarization in the chain structure formation of the single stranded mixture is prevented. This result may be explained with the prevention of the chain structure and reduction of the interaction between the magnetic particles in the single stranded structure which we explained before. According to us, the interaction between magnetic particles in the single stranded mixture mostly causes clusters.



(a)



(b)

Figure 6. Transmission spectrum of magnetic liquid that are without DNA, 20mg double- stranded DNA and single-stranded DNA under 100G magnetic field. a) P= 90 polarization, b) P=0 polarization

These observations can be explained by the interaction between DNA and NPs which prevents the chain formation. The study conducted by Byrne et al. (2004) examined the Fe3O4 NPs synthesized with the co-precipitation inside a solvent containing a single strand denature and double helix herring DNA and it was observed that the NPs were bonded to the DNA molecule with the formation of Fe-O-P bonds [17]. Transmission increase at both polarizations can be explained by formation of another structure of MNPs instead of chain structure. It is clear that the new structures of NPs open the light path and increase transmission due to the single stranded DNA in liquid.

5. Conclusions

In this study, MNPs were synthesized using the coprecipitation method with a mean particle size of 34 nm and a saturation magnetization of 62 nm. In order to obtain the measurements, the appropriate experimental set up was prepared and the MOLD effect was investigated for with single- stranded DNA, double stranded DNA and without DNA. The DNA-MNPs solution obtained was examined for the change in polarization of the mixture of single and double stranded DNA and magnetic fluid. The MOLD effect is shown. As the magnetic field increases, the chain structure is reinforced and the impact of linear dichroism is determined to increase.

When DNA is added to the reference fluid, it is seen that the effects related to polarization in the transmission of light is reduced. Together with the NPs in the fluid, the presence of DNA affects the chain formation mechanism of the MNPs. As the DNA amount increases, the interaction between the MNPs decreases.

Along with the measurements made with double helix DNA; single stranded DNA which is very important for the diagnosis of certain diseases, was examined to show the changes in the transmission of light under the magnetic field and related to polarization with the help of the same experimental apparatus. When the single stranded, double stranded DNA containing samples and the samples that do not contain DNA molecule are compared, it is revealed that the changes based on polarization in the double helix DNA-magnetic fluid structure show the formation of a chain structure. However, transmission of light related to polarization is lower compared to the fluid which does not contain DNA. Changes related to polarization is considerably low in the single stranded version.The obtained results will contribute to the development of studies with regard to the use of DNA-MNPs mixture as DNA sensor or biosensor.

6. References

[1] R. E. Hernandez , A. Baeza & M.V. Regi , “Smart Drug Delivery through DNA/Magnetic Nanoparticle Gates.”, *ACS Nano,* 5(2), pp.1259-1264. 2011.

[2] Ge, Jianping, et al. "Superparamagnetic magnetite colloidal nanocrystal clusters." *Angewandte Chemie International Edition* 46.23 ,4342-4345, 2007

[3] G.Jianping,Y.Hu, and Y.Yin.,"Highly tunable superparamagnetic colloidal photonic crystals." *Angewandte Chemie*, 119.39, 7572-7575, 2007

[4] S.Mostafa, et al. "Terahertz Faraday rotation in a magnetic liquid: High magneto-optical figure of merit and broadband operation in a ferrofluid." *Applied Physics Letters* 100.24 ,241107. 2012

[5] J., Min Sook, et al. "Methodological considerations of electron spin resonance spin trapping techniques for measuring reactive oxygen species generated from metal oxide nanomaterials." *Scientific reports* 6 (2016).

[6] M.T. Pooja ,V.Komal, A.D. Vida, & R.S. Shree, “Functionalized Gold Nanoparticles and Their Biomedical Applications.” *Nanomaterials*, s.31-63 ,2011.

[7] Melancon, M., & Lu, W. “Gold-Based Magneto-Optic Nanostructures: Challenges for In Vivo Applications in Cancer Diagnostics and Therapy.” *Mater Res. Bull*, s.415-421 ,2009

[8] Pershina, Alexandra Gennadievna, A. E. Sazonov, and Victor Dmitrievich Filimonov. "Magnetic nanoparticles–DNA interactions: design and applications of nanobiohybrid systems." *Russian Chemical Reviews*, 83.4, 299, 2014.

[9] Tiwari, A. P., S. J. Ghosh, and S. H. Pawar. "Biomedical applications based on magnetic nanoparticles: DNA interactions." *Analytical Methods* 7.24,10109-10120,2015

[10] S.Y. Park, H. Handa, and A. Sandhu. "Magneto-optical biosensing platform based on light scattering from self-assembled chains of functionalized rotating magnetic beads." *Nano letters* 10.2 ,446-451, 2009.

[11] Xu, M., and Ridler, P. J., , “Linear dichroism and birefringence effects in magnetic fluids”, *American Institute of Physics*. 82, 326, 1997

[12] Inaba, N., et al. "Magneto-optical absorption in infrared region for magnetic fluid thin film." *IEEE Transactions on Magnetics* 25.5 ,3866-3868. 1989

[13] Rousan, Akram A., Hassan M. El-Ghanem, and Nihad A. Yusuf. "Faraday rotation and chain formation in magnetic fluids." *IEEE Transactions on Magnetics* 25.4 ,3121-3124, 1989

[14] Ranjbar, Bijan, and Pooria Gill. "Circular Dichroism Techniques: Biomolecular and Nanostructural Analyses- A Review." *Chemical biology & drug design* 74.2 ,101-120,2009

[15] Köklü, S.Ö., Yariçi, İ., and Öztürk, Y., "Investigation of DNA-Magnetic Nanoparticle Interaction by Magneto-Optical Linear Dichroism ", *Biyomut16, 2016*
[16] Cao, S. W., Zhu, Y. J., & Chang, J. "Fe3O4 polyhedral nanoparticles with a high magnetization synthesized in mixed solvent ethylene glycol–water system." *New J. Chem*., s.1526–1530 ,2008
[17] Byrne S. J.,Corr S. A., Gun'ko Y. K., Kelly J. M., Brougham D. F., and Ghosh S.." Magnetic nanoparticle assemblies on denatured DNA show unusual magnetic relaxivity and Potential applications for MRI", *Chem. Commun*., 2560, 2004

Selma KÖKLÜ was born in Alaşehir district of Manisa. Selma has enrolled at university in 2007 and graduated with an honour degree-3rd of department from the department of physics at Izmir Institute of Technology in 2012. In the same years, she studied extra course in electrical and electronics engineering department. She did willing internship at Atomic Energy Institute of Turkey in Istanbul. In 2013 Selma applied to Ege University for master degree in Biomedical Technologies and completed her master's degree in 2016. In the master thesis she investigated the active attachment of DNA and magnetic nanoparticles and magneto-optical effects of single-stranded and double-stranded DNA patterns by polarization optics.

İsmail YARİÇİ was born in Ağrı, Turkey. He received the B.E., M.E., in Electrical & Electronics Engineering, Izmir Institute of Technology (IZTECH), Turkey, in 2012, and Ege University, Turkey, in 2015, respectively. He is currently a Ph.D. student in Ege University.In 2012 he joined the department of Electrical and Electronics Engineering, Ege University, as a research assistant and he is currently working there. His current research interest include magneto-optic measurement systems, biomedical optic and electromagnetic teori.

Yavuz Öztürk was born in Kayseri , Turkey. Yavuz Öztürk received his MSc (2004) and PhD degrees (2010) in Electrics and Electronics Engineering from Ege University (İzmir, Turkey). He was a postdoctoral fellow at INRS-EMT between 2010-2012. Since 2012, he is a assistant proffesor at Ege University. He studied the fabrication and characterization of garnet type magneto-optical films, magneto-optical measurement systems, magnetic material properties and spin dynamics.

A Fully Automatic Novel Method to Determine QT Interval Based on Continuous Wavelet Transform

Cüneyt YILMAZ[1], Mehmet İŞCAN [1], Abdurrahman YILMAZ[1]

[1]Mechatronics Engineering Department, Yıldız Technical University, İstanbul, Turkey
abyilmaz@yildiz.edu.tr,miscan@yildiz.edu.tr,cuneyt@yildiz.edu.tr

Abstract: *Nowadays, automatic recognition algorithm is being frequently utilized to extract the information concerning cardiac abnormalities. In this study, a fully automatic novel method based on the continuous wavelet transform (CWT) was developed for QT intervals in various ECG signals. Especially, the determination of T-wave end is the paramount problem to be solved.The developed method was performed to find the beginning of QRS complexes and the end of T-wave. The proposed algorithm was tested on MIT-BIH-NSR database given by QT database, then, it yielded the scores 15.17 milliseconds and root-mean-square error of 17.19 milliseconds at silver standard, 19.22 milliseconds and 20.22 milliseconds at gold standard, respectively. In conclusion, the proposed algorithm is a fully automatic method to attain a high performance in the calculation of QT intervals at various ECG signals.*

Keywords: *ECG Signal, QT interval, ECG signal classification, Pan-Tompkins Algorithm, Continuous Wavelet Transform*

1. Introduction

Electrocardiography(ECG) is represented as the cardiac activity reflecting its electrical activity in heart muscles. Since the invention of the Einthoven rules, it has been used as a diagnostic feature to identify electrical propagation of cardiac muscles, which means that it has valuable information regarding cardiac situation of a patient. Generally, the cardiac activity is separated into two parts, depolarization and repolarization indicating contraction and relaxation of heart muscle. Therefore, some sort of abnormal electrical propagation are reflected in ECG, then, one can extract vital features in cardiac level of patients.

Namely, ECG consists of two waves (P and T) and one complex (QRS) which are illustrated in Fig.1.

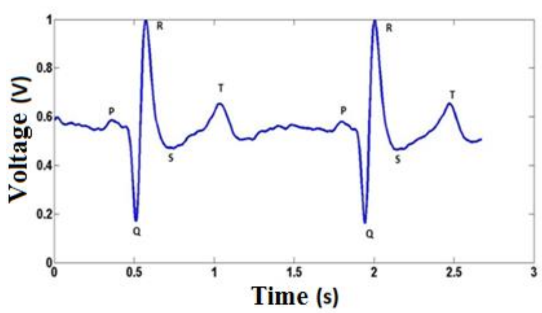


Fig. 1. ECG Signal

P wave denotes the contraction of atrial muscle pumping the blood into the ventricles. QRS complex represents the stimulation of right and left ventricles. Right and left ventricles perform to pumps venous blood into the pulmonary and fresh blood through the artery. All of the works to be done are automatically achieved which means that the heart is an autonomous system. Synchronization is provided by the electrical stimulation within specific intevals(QT inteval, PR segment). As a results, any irregularities including mechanical and electrical abnormalities in heart synchronization causes to fail the heart muscle leading to sudden death [1].

QT interval, one of the most important duration, is defined as the time lapse between the beginning of the depolarization and the end of the repolarization in heart ventricles. There is a strong relationship between QT intervals and several heart diseases, in which it is utilized to diagnose some of the abnormalities [2].

Before the advancement of computer programming paving a way for easily solving difficult problems, Experts and Cardiologists decided that electrical activities of heart had whether abnormalities or not , by examining the ECG signals in terms of duration and amplitude features, which were literally manually measured.

However, the performance of the manual QT measurement was reduced because of long records, fatigue in reading, individual evaluation and ECG device feature [2]. In addition, only 25% of cardiologists who had no experience in electrophysiology could assign correct marks of QT interval locations[3]. In long records, manual measurements had a problem with selecting representative beats indicating instantaneous cardiac situations due to the non-stationary feature of ECG [4].

Received on: 31.01.2017
Accepted on: 13.03.2017

For these reasons, automatic algorithms have been developed to determine the beginning of the QRS complex and the end of the T-wave.

Numerical methods based on the derivative feature were performed to measure QT intervals [2,6]. In despite of reduction in the calculation of the cost, the performance of QT measurement was affected by T wave amplitude, isoelectric baseline changes, and false detections. Also, the T-wave slope and the isoelectric level are required to be applied these algorithms. In addition, intra-beat variations have a negative effect on detecting boundary points in the ECG pattern.

Some of the methods related to formulation of ECG patterns were established to analyze the start and end points of ECG waves [7,8]. Although, these algorithms can be effectively applied to classify basic waves such as QRS and T, they had a poor score for the determination of boundary points due to the vexing problem which was defined as trying to use a deterministic method applying a non-deterministic problems [9].

In order to process non-deterministic ECG signals, neural network applications were developed [10-12]. Neural networks can be efficiently achieved to measure QT intervals regardless of frequency characteristics of ECG, after training them. However, this process requires too long training data, and is not an automated method.

There are some methods which are independent of ECG level and implemented in the frequency domain [13]. It is noteworthy that the measurement performance decreases with significantly removing waveforms,even the algorithms are fully-automated. On the other hand, these algorithms cannot be achieved to apply on the signals obtained from different channels of ECG because T wave and QRS complexes orientations change.

Due to all reasons explained above, in this study, a novel full-automatic algorithm based on continuous wavelet transform (CWT) was proposed to recognize the beginning of depolarization and the end of repolarization in heart ventricles. In the presented algorithm, the ECG signals were preprocessed by filtering and flattening. Then, the ECG boundaries are marked by using CWT. The algorithm performance was evaluated in terms of gold and silver standards which are the comparisons with manual annotations and comparison of standard automatic algorithms, respectively.

2. Method

QT intervals in ECG records were computed in this study and the evaluation was performed by comparing achieved and expected results (annotations) . QT interval detection and evaluation steps of the algorithm will be investigated in the subsections.

2.1. QT Interval Detection

2.1.1. Preprocessing

An ECG record is a quite rough signal due to ambient and measurement noises, undesirable effects on the physical conditions such as patient’s breathing, the activities of other organs. To eliminate such conditions, the ECG records were filtered by a band pass FIR (Finite Impulse Response) filter with 0.5 Hz and 40 Hz cut-off frequencies since the electrical signals produced by the human heart during the pulse were in this frequency band [6]. The mathematical description of the filter is

$$y(n)=\sum_{k=0}^{N} c_k \cdot x(n-k) \qquad (2.1)$$

where c_k is a constant scaling parameter, x is to be filtered the ECG record, y is the filter output, n is the current time step and $N=330$, the filter order. The filter order is selected such a large value which is an appropriate number containing one electrical cycle of the heart in order to take the desired frequency window of the records exactly. On the other hand, the computational time for filtering is increased, which means that such a selection was not suitable for real time applications.

Coiflet-1 type continuous wavelet transform was applied to the filtered ECG record. Type 1 represents that the used wavelets are orthogonal wavelets. Scaling function level of the wavelet is selected as 9 with respect to the frequency band of T wave end pattern [14]. The resulting signal was used for QT interval detection analysis as explained in the following sections..

2.1.2 QRS Complex Starting Point Determination

Firstly, $QT\ Pattern$ signal which will be used for determination of QRS complex starting and T wave end points was defined as

$$QT\ Pattern=(CWT\ Output)^2 \qquad (2.2)$$

where $CWT\ Output$ is the signal obtained after CWT analysis. This mathematical trick was used to make higher and lower amplitude parts of the signal more apparent. The amplitude of the $QT\ Pattern$ is high in the vicinity of the R peaks, therefore to be analyzed signal for QRS complex starting point $QRS\ Pattern$ is defined with a piecewise function as

$$QRS\ Pattern(i)=\begin{cases}0 & ,QT\ Pattern(i)<A\\ QT\ Pattern(i), & otherwise\end{cases} \qquad (2.3)$$

where A is equal to three times mean values of $QT\ Pattern$, i is an integer number from 1 to ECG record length. This comparison level to produce $QRS\ Pattern$ was found with a heuristic approach.

In $QRS\ Pattern$, all the peaks are possible R peaks of the evaluated ECG record, and the highest amplitude peaks in the defined scan range are R peaks of the ECG record. R peak scan range was selected as ±160 milliseconds because of the largest healty QRS duration[15].

By using R peak locations information, the QRS complex starting points can be determined. However, the directions of the R peaks in the ECG record should be considered. The R peak directions may be positive or negative based on its record lead. Therefore, the piecewise function (2.4) considering R peak directions was used to mark QRS complex starting point as

$$QRS_{Start}(i) = \begin{cases} Q_{Min}(i) - 8\ ms, & R_{Peak}(i) \text{ is positive} \\ Q_{Max}(i) - 8\ ms, & R_{Peak}(i) \text{ is negative} \end{cases} \tag{2.4}$$

where QRS_{Start} is QRS complex starting points, i is an integer number from 1 to ECG record length, Q_{Min} and Q_{Max} are minimum and maximum points around left side of R peaks respectively and R_{Peak} is investigated R peak of the ECG record. It is obvious that minimum or maximum levels were not directly marked as QRS complex starting points. Eight millisecond subtraction was used to catch points at isoelectric level. The first phase of the QT interval determination was finalized by labeling the starting points of QRS complexes.

2.1.3 T Wave End Point Determination

In the previous section, it has been declared that direction of R peaks could be positive or negative. The same situation is also valid for T-wave of the ECG records. For correctly positioning the T-wave end on the ECG, the directions of the T-waves should be identified as positive or negative. For this purpose, the method shown in Fig.2 was used. Firstly, RR interval for each beat shown in Fig.1 was calculated by the times between consecutive R peaks. Secondly, the highest and lowest amplitude points were marked as possible T-wave peaks of the ECG record by scanning the next half of the RR interval from the corresponding R peak. Finally, the direction of T-wave peaks whether positive or negative was determined by comparing the area under the ECG curve and the shaded area as shown in Fig.1. As a result of this decision mechanism, the highest or lowest possible T-wave peak points previously marked were assigned as T-wave peaks.

The amplitude of the signal achieved after CWT analysis is relatively low around T-waves while these parts of the signal were critical for the determination of T-wave end points. To focus the desired signal windows, piecewise function expressed in (2.5) was used.

$$T\ Pattern(i) = \begin{cases} 0 & , QT\ Pattern(i) > 0.05 \\ QT\ Pattern(i), & otherwise \end{cases} \tag{2.5}$$

where $T\ Pattern$ represents the signal to be used for detection of T-wave ends and i represents an integer from 1 to ECG record length. The threshold value, 0.05, was selected to remove the all frequency peaks produced by continuous wavelet transform at T wave except the end of the repolarization, in ventricles, T wave end. Normally, T wave frequency peaks generated by the continuous wavelet transform establish several critical points indicating the turning place on the ECG signal. One of them, the T wave end possesses the three crosspoints which are descenting, ascending and steady line, respectively. The threshold values clarify these lines which can be detectable. Achieved $T\ Pattern$ signal was normalized before using it, because it contains signal parts with only weak amplitudes. Even if the high-amplitude parts are cleaned, the $T\ Pattern$ may still contain sections related to the T-wave ends and that may cause faulty evaluation. The Clear Area Region (CAR) shown in Fig.3 is a window starting from 20% of RR interval in the rear of R peaks up to the previously determined T-wave peaks. The CAR of each beat was suppressed in $T\ Pattern$ to eliminate the parts irrelevant to the end of the T-waves. The rest of the signal was used to detection of T-wave end positions.

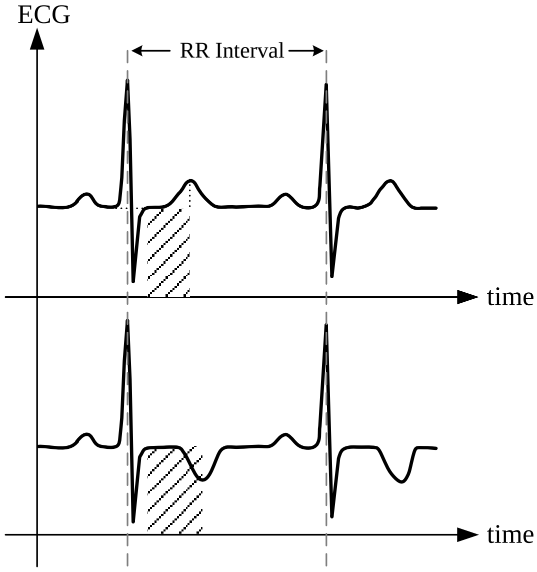


Figure 2: Positive or Negative T-wave Peak Determination

In the last step of QT interval determination, the algorithm focused the right hand sides of the previously determined T-wave peaks. Therefore, the peaks following the T-wave peaks of the cleared signal were utilized. The decision mechanism shown in Fig.4 was executed to mark T-wave ends of each beat in ECG record. For previously found each T-wave peak in ECG record, the peaks of the cleared were scanned in T-wave scan range. When both the 2nd and 3rd peaks are available in the cleared , the T-wave end position of the beat was assigned as the average position of the 2nd and 3rd peaks. On the other hand, the T-wave end position of the beat was assigned as the position of 2nd peak if it is available, but the 3rd one is not. Otherwise, it was assumed that the T-wave end of the beat is not detected.

QT intervals of the beats in ECG record was calculated by using the following function

$$QT\ Interval(i) = T\text{-}wave\ end(i) - QRS\ start(i) \tag{2.6}$$

where $QT\ Interval(i)$, $T\text{-}wave\ end(i)$ and $QRS\ start(i)$ are QT interval value, T-wave end position and QRS complex start position of the ith beat, respectively.

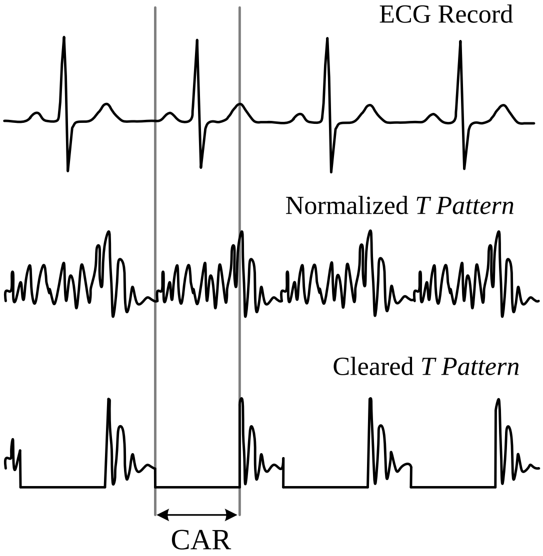


Figure 3: Acquiring the signal to be used for calculation of the T-wave end points

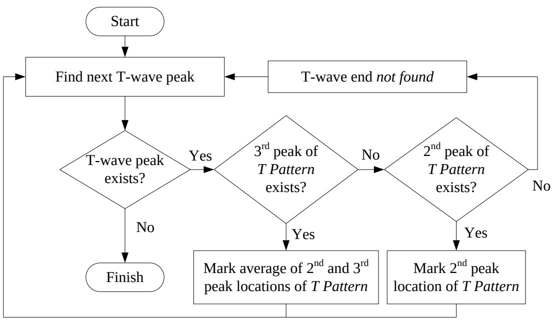


Figure 4:T-wave end position determination flowchart

2.2 Evaluation Procedure

QT interval values determined by the proposed algorithm in this study and received from the annotations were compared in terms of mean error and standard deviation error. After that the score of the algorithm for evaluated ECG record database was calculated as mean squared error (MSE) of the previously found QT interval mean errors for each evaluated ECG record.

3. Experiment

The experiment was performed by using MIT-DB-NSR and MIT-BIH Long-Term ECG signal records in QT-database [16]. In MIT-DB-NSR and MIT-BIH Long-Term ECG database, there are 15 records annotated by the experts and standard automatic algorithms. All records are sampled at 250 Hz, and are normalized to remove changes of amplitude level in ECG. In addition, the presented algorithms was carried out on the whole QT database in order to compherensively compare the other studies published in later.

The algorithm was performed on the database given above. In the first place, the CWT is applied on the ECG signal in order to get the fiducial points such as Q and R points. The Figure 5 illustrates how the algorithm works..

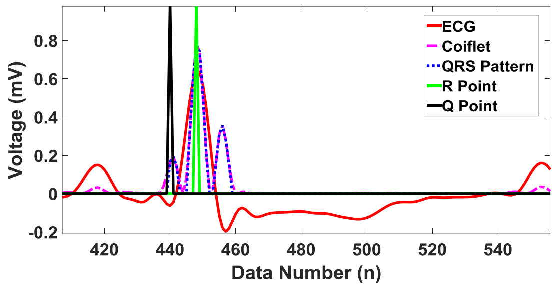


Fig. 5. Q and R point detection

In Figure 5, X and Y axis represent the data number and the voltage value in ECG. The coiflet pattern obtained by CWT was transformed into QRS pattern shown by three peaks on ECG. The first and second marks are assigned as Q and R point, respectively.

Then the algorithm should find where T peak occurs before obtaining the T wave end location. Figure 6 schematically shows the algorithm's annotation.

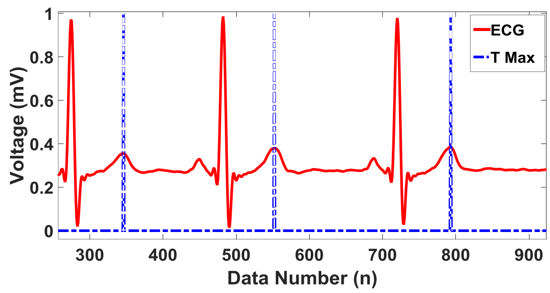


Fig. 6. T wave maximum marking

In Figure 6, the T wave maximum is obtained by searching maximum on refractory period produced by computing RR interval at each cycle.

Finally, T wave end point is marked by evaluating Coiflet pattern in terms of sequential peaks. Figure 7 clearly explains the marking process on ECG signal.

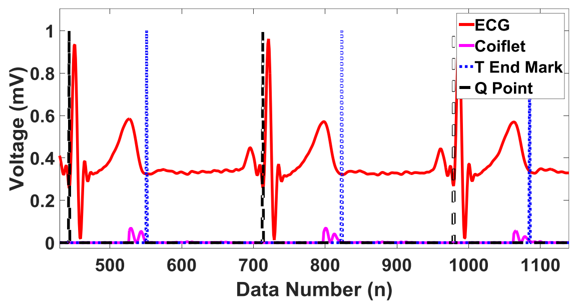


Fig. 7. T Wave End Detection

In Figure 7, Coiflet patterns are only placed at right hand side of T maximum because of erased peaks before T maximum happens. It can be shown that T wave end is generally found between the third and the second peaks of Coiflet patterns. Actually, Coiflet patterns were established by nine individual frequency regions, which means that each of the frequency regions had a specific sub-signal on ECG illustrating the feature of depolarization and repolarization of the time-related action potential. On the graph (Figure 7), it could be understood that each coiflet peak happens when the sign of derivative of ECG wave changes. This led us to estimate where the repolarization was finished and started. The detailed information has been given in the Method section of this article.

The proposed algorithm applied on the all databases had the scores illustrated in Table 1.

	Manual	Automatic
Mean Error	19.22 ms	15.17 ms
RMS Error	20.22 ms	17.19 ms

Table 1. Test Results

Table 1 shows manual and automatic measurement errors in the tests. According to these results, the manual tests had poor score than automatic tests because manual annotation was only chosen at representative beats which characterize individual cardiac status of patients and had the same patterns. Also, electrophysiologists almost marked the points indicating T wave end before the algorithm one [17]. This led to a great difference in mean error representing how much the algorithm was effective.

The automatic test was performed on all ECG records including broken patterns by using ECGPUWAVE method which had a challenging performance. These results are deeply discussed in the next section.

4. Discussion

In this study, a novel fully automated method was proposed to analyze QT interval in ECG signals. The algorithm achieved QRS complex recognition, finding T-wave peaks and end points in all records. All ECG records were obtained from QT-database which were annotated by the experts and automatic algorithm. However, QT-database contained five different abnormality groups which were not included in this study.

European ST-T Databases included many different patients with myocardial infraction, cardiac artery diseases and hypertension causing ST elevation and rippling isoelectric baseline in ECG. Generally, V4-V5 channels of ECG were used to collect cardiac signals. Due to these reasons, CWT produced a lot of wrong results in the tests. There was no consensus where the T-wave peaks were even in a standard automatic algorithm. Besides, the respiration which was a primary cause of a low frequency noise was not stable in these patient groups. The collision between respiration and T wave end gave many negative results.

We could not include the results from MIT-BIH Supraventricular arrhythmia database, because of their different sampling frequencies. Although the transformation in the sampling frequency was done, there was only one annotation marked by only one cardiologists. So, this database was not included in this study due to much more requirement in control of the proposed algorithm.

Records from Sudden Death patients in BIH database had different T wave morphology, positive negative T-wave. The proposed algorithm was utilized at only positive or only negative T-wave morphology. Nevertheless, these databases results had poor results obtained by standard automatic annotations. In addition, T-wave morphology was not easily discriminated even by observation. As the same way, MIT-BIH ST change database contains records with the high heart rate, so, the ripple in the signal is too high level. MIT-BIH Arrhythmia database was sampled at a different frequency, and contained a different T-wave morphology which our algorithm could not be applied. Additionally, these databases consisted of many premature ventricular contraction (PVC) records which meant that RR interval mostly changed causing to impact on QT interval variation, and the proposed algorithm performance too.

On the other hand, the proposed algorithm did not require the determination of isoelectric level. The algorithm searched specific frequency energy which included T-wave end. The algorithm was not affected by the amplitude level, but frequency changes. It discriminated T-wave positive or negative peaks in MIT-BIH-NSR and Long-term records. Additionally, the decision mechanism related to finding T-wave positive-negative was mostly affected by Q points deflection, because some of the changes such as ST-T elevation and PVC shifted the level of Q points. Mostly, the other database errors were also caused because of that.

Up to now, several studies have been performed to extract T wave end information from QT database [2,10,18,19,20,21,22,23]. The related results are given in Table 2.

	Mean Error	Standard Deviation
Laguna et. al.	18.86	29.79
Cesari et. al.	-0.6	22.3
Leon et. al.	-0.12	15.06
Madeiro et. al.	2.8	15.3
Manriquez et. al.	0.57	22.81
Martinez et. al.	-1.6	18.1
Vila et. al.	-2.7	28
Zhang et. al.	0.31	17.43
Present Work (only MIT-NSR)	0.92	8.72
Present Work (whole QT database)	1.01	12.44

Table 2. Comparison between The Published Algorithm

Table 2 shows the important information about the automatic algorithm. Laguna et. al. [18] used threshold detector in order to make a simple and effective algorithm on Holter device. The results could not efficiently be performed because of the real time application. Manriquez et. al. [21] applied likelihood ratio to detect T wave end, however, their study used multiple ECG signals. Some of the methods were based on wavelet approach [19.22]. In spite of high performance on T wave end location, there were no standard measurement results in order to compare that their algorithm were clinically acceptable or not. Generally, the derivative methods were simple and effective in non-noisy conditions[2,18]. However, there were too much operations to process the ECG signals, such as requirement of isoelectric level, T wave slope and maximum points changes. The other algorithms reported were implemented on neural network and mathematical modeling. So, the training and comparing phase are highly expensive, and also not fully automated algorithm.

The proposed algorithm was tested on two conditions: only for MIT-BIH-NSR and MIT-Long Term records, and the whole QT database. The performance of the algorithm was quite sufficient in MIT-BIH-NSR and Long-term records due to proper waveform characteristics of ECG signals. However, the measurement performance in evaluating the whole QT database was degraded by some erroneous properties such as different T wave form, absent or corrupted waveform, excessive noise, etc. In spite of all disadvantages mentioned earlier, the whole QT database evaluation results attained high performance classification.

Up to now**,** the standard algorithms for the detection of QT interval were extensively studied. However, the standard methods are not useful to produce clear information about the cardiac situation of the patient under some dangerous conditions. Some of the diseases, especially in Long QT syndrome causing sudden cardiac death, the standard deviation of the measurement error for detecting the illness is so low which the existing algorithms does not meet this criteria. The presented algorithm can be performed with the standard deviation of 8.72 milliseconds while comparing the expert results. It is concluded that the suggested algorithm provide a better information in comparison with the previous studies.

On the other hand, all studies given by references were differently evaluated in terms of mean and standard deviation error. The calculation of the mean error was done by summing all of the time errors between annotation and the algorithm regardless of their sign. This led to zero mean error if the number of positive mean error equal to negative ones. For this reason, the proposed algorithm was evaluated by using the standard measurement given in [17]. According to this study, absolute mean error and standard deviation error could be acceptable as 15 milliseconds and 20 milliseconds, respectively. As a result, the proposed algorithm could be considered as a diagnostic tool in clinical applications.

Generally, manual tests had poor scores due to lack of representative beat selections. Additionally, automatic algorithm results may be improved by adding representative beat controls in the classification process.

5. Conclusions

In this study, a fully automated technique based on CWT was proposed to measure QT intervals in various ECG signal patterns. Especially, marking T-wave end is the most difficult problem. In this study, CWT was firstly applied on the ECG signals to detect Q position, and then T-wave end location was obtained by the algorithm. The proposed algorithm had achievements both at 15.17 ms mean error and 17.19 ms RMS error in the automatic test and at 19.22 ms mean error and 20.22 ms RMS error in the manual test.

The proposed algorithm does not need any training data, or manual intervention, and also was not influenced by isoelectric level changes and intra-beat variations. However, it is only performed on only positive or only negative T-wave morphologies. Moreover, All QT-databases were not included in the tests because of the existence of different abnormalities and the restrictions on the present algorithm.

In the future work, RR interval changes and selective representative will be added to the algorithm. Additionally, six different T-wave morphologies will be examined to update the algorithm.

In conclusion, the results showed that the proposed algorithm can be effectively used to attain a high performance in the determination of QT intervals automatically.

6. References

[1] Goldberg, R. J., Bengtson, J., Chen, Z., Anderson, K. M., Locati, E., & Levy, D.,"Duration of the QT interval and total and cardiovascular mortality in healthy persons (The Framingham Heart Study experience)".,The American journal of cardiology, 67(1), 55-58.,1991

[2] Zhang, Q., Manriquez, A. I., Médigue, C., Papelier, Y., & Sorine, M.,"An algorithm for robust and efficient location of T-wave ends in electrocardiograms",*IEEE Transactions on Biomedical Engineering*, 53(12), 2544-2552,2006.

[3] Goldenberg, I. L. A. N., Moss, A. J., & Zareba, W. QT interval: how to measure it and what is" normal". *Journal of cardiovascular electrophysiology*, *17*(3), 333-336, (2006).

[4] Helfenbein, E. D., Zhou, S. H., Lindauer, J. M., Field, D. Q., Gregg, R. E., Wang, J. J., ... & Michaud, F. P. An algorithm for continuous real-time QT interval monitoring. *Journal of electrocardiology*, *39*(4), S123-S127, (2006).

[5] Murray, A., McLaughlin, N. B., Bourke, J. P., Doig, J. C., Furniss, S. S., & Campbell, R. W.,"Errors in manual measurement of QT intervals",*British Heart Journal*, 71(4), 386-390, 1994.

[6] McLaughlin, N. B., Campbell, R. W. F., & Murray, A.,"Accuracy of automatic QT measurement techniques". *Computers in Cardiology*, Proceedings. 863-866. IEEE, 1993.

[7] Hayn, D., Kollmann, A., & Schreier, G.,"Automated QT interval measurement from multilead ECG signals",*Computers in Cardiology*, 381-384, IEEE, 2006.

[8] Clifford, G. D., & Villarroel, M. C.,"Model-based determination of QT intervals",*Computers in Cardiology*, 357-360, IEEE, 2006.

[9] Xue, Q., & Reddy, S. Algorithms for computerized QT analysis. *Journal of electrocardiology*, *30*, 181-186. (1998).

[10] Alexander A. Suarez Leon et al, "Neural Network Approach for T wave End Detection: A comparison of Architectures", Comput. Cardiol., vol. 42, pp. 589-592, 2015.

[11] Mehmet İşcan; Cüneyt Yilmaz; Faruk Yiğit, " T-wave end pattern classification based on Gaussian mixture model", DOI: 10.1109/SIU.2016.7496149, Pages 1953 - 1956, 2016.

[12] Mehmet İşcan, Faruk Yiğit, Cüneyt Yilmaz," Heartbeat Pattern Classification Algorithm Based on Gaussian Mixture Model", Inter. Symp. on Med. Meas. and App., Italy,2016.

[13] Chesnokov, Y. C., Nerukh, D., & Glen, R. C.,"Individually adaptable automatic QT detector". *Computers in Cardiology*,337-340, IEEE, 2006.

[14] Li, C., Zheng, C., & Tai, C. Detection of ECG characteristic points using wavelet transforms. *IEEE Transactions on biomedical Engineering*, *42*(1), 21-28., (1995).

[15] Chudacek, V., Huptych, M., Novak, D., & Lhotska, L.. Automatic QT interval measurement using rule-based gradient method. In *Computers in Cardiology, 2006* (pp. 349-351). IEEE, (2006, September)

[16] Laguna P. R. G. Mark, Goldberger A., Moody G. B., "A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG", *Comput.in Card*., 1997.

[17] Willems, J. L., Abreu-Lima, C., Arnaud, P., van Bemmel, J. H., Brohet, C., Degani, R., ... & Machado, H. The diagnostic performance of computer programs for the interpretation of electrocardiograms. *New England Journal of Medicine*, *325*(25), 1767-1773., (1991).

[18] Jané, R., Blasi, A., García, J., & Laguna, P.. Evaluation of an automatic threshold based detector of waveform limits in Holter ECG with the QT database. In *Computers in Cardiology 1997* (pp. 295-298). IEEE., (1997, September)

[19] Cesari, M., Mehlsen, J., Mehlsen, A. B., & Sorensen, H. B. D. Application of a new robust ECG T-wave delineation algorithm for the evaluation of the autonomic innervation of the myocardium. In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the* (pp. 3801-3804). IEEE, (2016, August).

[20] Madeiro, J. P., Nicolson, W. B., Cortez, P. C., Marques, J. A., Vázquez-Seisdedos, C. R., Elangovan, N., ... & Schlindwein, F. S. New approach for T-wave peak detection and T-wave end location in 12-lead paced ECG signals based on a mathematical model. *Medical engineering & physics*, *35*(8), 1105-1115., (2013).

[21] Manriquez, A. I., Zhang, Q., Medigue, C., Papelier, Y., & Sorine, M. Multi-lead T wave end detection based on statistical hypothesis testing. *IFAC Proceedings Volumes*, *39*(18), 93-98., (2006).

[22] Martínez, J. P., Almeida, R., Olmos, S., Rocha, A. P., & Laguna, P. A wavelet-based ECG delineator: evaluation on standard databases. *IEEE Transactions on biomedical engineering*, *51*(4), 570-581., (2004).

[23] Vila, J. A., Gang, Y., Presedo, J. M. R., Fernández-Delgado, M., Barro, S., & Malik, M. A new approach for TU complex characterization. *IEEE Transactions on Biomedical Engineering*, *47*(6), 764-772., (2000).

Note:

Abdurrahman Yılmaz was born in Muğla. He finished science high school in Niğde. In order make true his dream about to be an engineer; he got into Istanbul Technical University Electronics Engineering Program. At the beginning of second class, he also enrolled Control & Automation Engineering at the same university by means of double major program. Before graduation, he has started to work for "ASELSAN" which is the biggest defense industry company in Turkey for 7 months. In order to continue academic life, he started to work and still working as a research assistant at Yıldız Technical University. Now, he is a master student on Mechatronics Engineering Department at Yıldız Technical University and he has been working in areas such as the interpretation of ECG signals and walking analysis and control of the robotic systems.

Mehmet İşcan was born in Malkara, Tekirdağ. He got bachelor and master degree from Yıldız Technical University at Mechatronic Engineering. He was accepted as a candidate student of Phd program at the same department. He has also been working at Mechatronic Engineering Department in Yıldız Technical University since 2014. On the other hand, he is a co-founder of a Pulse&More Corporation, which is a biomedical company developing and producing cardiac testing device, Holter,etc. He is also interested in LVAD design, neural network, and detection of cardiac abnormalities of heart muscle.

Cüneyt Yılmaz was born in Beykoz, Istanbul. He studied Mechanical Engineering in Istanbul Technical University. After he got his masters degree, he went to the United States for his PhD education. After he got his PhD degree, he worked in University of Texas Southwestern Medical Center's Pulmonary laboratories as a senior research engineer, research faculty and assistant professor. After 17 year experience in the United States, he returned back to his native country, Turkey. He has been a faculty member of the Mechatronics Engineering Department of Yıldız Technical University Mechanical Faculty for about 3 years.

ASSESSMENT OF SIMILARITIES BETWEEN LIVER IMAGES TO EACH OTHER USING SCALING, ROTATION AND TRANSLATION GEOMETRICAL OPERATIONS

Tuğba Palabaş Tapkın[1], Onur Osman[2], Tuncer Ergin[3], Uygar Teomete[4], Özgür Dandin[5], Nizamettin Aydın[6]

[1]Department of Computer Engineering, Yildiz Technical University, Istanbul, Turkey
[2]Department of Electrical Electronic Engineering, Istanbul Arel University, Istanbul, Turkey
[3]Radiology Department, GATA, Ankara, Turkey
[4]Department of Radiology, Miami Miller Faculty of Medicine, Florida, A.B.D
[5]General Surgery Service, Bursa Military Hospital, Bursa, Turkey
[6]Department of Computer Engineering, Yildiz Technical University, Istanbul, Turkey
tugba.palabas@gmail.com, onurosman@arel.edu.tr, tuncerergin@yahoo.com, uygarteomete@yahoo.com, dandinozgur@gmail.com, nizamettin@ce.yildiz.edu.tr

Abstract: *In this study, similarity rates of the liver images which are obtained from different peoples are determined using 3D geometric transformation methods. The similarity is evaluated based on the numerical comparisons and visual results. 10 intact liver images which are drawn by the radiologists are used. Three geometric transformation methods scaling, rotating, and translating are consecutively applied to the liver images. All images are used both as atlas and as test images. The Dice coefficient values are calculated to show the similarity of each test image to atlas. The scaling, rotating, and translating amounts of the image are retained for the atlas which the similarity rate is highest. The liver images of different persons are similar to each other at an average rate of 67 ∓0.09 % according to Dice coefficient values which express the similarity. This study is presented as a step to prepare atlas database for segmentation of the injured liver.*

Keywords: *Geometric operetions, liver segmentation, Dice coefficient.*

1. Introduction

In order to determine the immediate operation necessities and to make the pre-diagnosis of the pathologic findings to the traumatic patients which are brought to emergency rooms is very important for accelerating patient management.

The images that make up the CT (Computed tomography) visualising method will be the most accurate detection method. The first and the most important step is the segmentation of the liver who has similar grey level values with the neighboring organs in an accurate manner in order to re-building and the determination of the injury level. The determination of the liver borders by hand needs extra expertise and is a time-consuming job because of the section amount. For this reason, automatic segmentation becomes important step in biomedical image processing for computer aided diagnosis. Lots of methods have been proposed in order to realise computer aided diagnosis of intraabdominal organs. In Figure 1, automatic segmentation [1] of liver and spleen images have been utilized by the aid of probabilistic atlas method. 257 BT images have been compared with liver and spleen images by the aid of probabilistic atlases and the success of the proposed method has been evaluated by examining 10 CT images in which the borders have determined by hand. For liver and spleen , Dice coefficient of 96.2% and 95.2%, Tanimato indexes of 92.7% and 91%, volume estimation error of 2.2 and 3.3, peak estimation error of 2.8 and 1.7, RMS error of 2.3 mm and 1.1 mm and mean surface distance have been obtained as 1.2 mm and 0.7 mm. It has been stated that the methodology can be utilised in the routine analysis of normal and enlarged liver/spleen images' segmentation. In the second reference, a study has been presented about the multiplanat fast marching method which is involved with grey level automatic segmentation. 60 CT images have been used in order to test the segmentation success. For liver segmentation, an average of 94% and for spleen segmentation, an average of 93% volume overlap values have been determined as a method which is a general segmentation method to obtain anatomic knowledge which is not critical. In the third reference, a probabilistic model has been proposed for multiple organ segmentation. The method is aiming to calculate the

Received on: 31.01.2017
Accepted on: 13.03.2017

probabilistic equivalents of voxel values. In order to reduce the calculation complexity and to obtain the best figure variations, for the optimization of principle component analysis and forecasting results, iterative conditional mod-expectation maximization methods have been utilised. For livers and kidneys, 72 training sets and 40 testing sets were utilized and metrics were obtained. With numerical and visual results, it has been stated that the method is performing roughly for the abdominal multiple organ segmentation. In reference 4, a strategic combination for active appearance model, live wire and graph cuts for 3 dimensional organ segmentation have been proposed. The method which is consisting of model formation, object recognition and qualification steps has been tested on clinical CT data set for liver, kidney and spleen segmentation. The metrics which is related to the proposed method has been compared with different methods in order to express segmentation success. In reference 5, for the computer aided diagnosis and laparoscopic surgery help for the intraabdominal organs, a roughly automatic segmentation method which depends on hierarchical atlas records have been proposed. Manually segmented 150 CT image database have been utilised in order to test the success of the method. Consecutively, similarity ratios of 94-93-70-92% values have been obtained for livers, kidneys and pancreases. In reference 6, 4 dimensional CT data has been utilised for the segmentation of intraabdominal organs by utilising anatomic and physiological properties of computer aided diagnosis applications and a similarity ration of 90.5% has been obtained. By applying algorithm steps, the obtained numerical and visual results have been evaluated and the effect of outlook, shape and location data has been expressed. Methods depending on shape and location information have been proposed in reference 7, for injured liver and in reference 8, for injured liver and spleen segmentation. In computer aided diagnosis applications, for intraabdominal organ segmentation, the shape, volume and location information has been utilised often. The damages and density differences that occur in the organs from traumas effect the correctness of these information in a negative manner and change the borders of the segmentation. Selver et al. have developed a 3 dimensional segmentation method for the evaluation of donors before the transplantation by utilising CT images. It has been proposed that, the method which has been developed by utilising artificial neural networks, is suitable for the clinical applications. In CT images, for automatic multiple organ segmentation, an atlas based method has been proposed and presented by Wolz et al. Çınar and Durkaya have proposed an approach for the picking out of the liver image that depends on the expansion principal by using MR images. Every segment has been evaluated separately , the borders of the liver heas been determined and a 3 dimensional segmented liver image has been obtained. Campadelli et al. have proposed a method for the spleen and kidney segmentation that depends on approach and rule based system method.

In this study, the borders of liver for different people's CT images have been determined manually and the anatomical and physiological similarities of the liver in various people has been expressed by utilising metrics. In this manner, the studies that will be carried out on the segmentation of the injured liver images' segmentation will have similar border values and by this way, the success in the segmentation will be improved.

2. Material and Method

2.1. Material

In this study, the CT images of 10 patients has been utilised. The images for abdomen trauma have been obtained by using standard protocol. All CT images in portal venous phase were 3 mm segmented and was composed of 84 and 164 various segments.

In this study, the borders of the liver images that have been determined by radiologists (T.E. and U.T.) and drawn by hand were utilised.

2.2. Method

The intraabdominal healthy organs have similar anatomic and physiological properties in different people. Geometric conversion methods have been utilised in order to demonstrate this case numerically. Geometrically, where the value of the pixel will be moved to is determined. The coordinate transformation of the image pixels or the interpolation of the pixel values can be applied as conversion methods. Image scaling (magnification, reduction), mirroring, rotating, cropping, shifting methods can be utilised to change image border in any axis.

In this study, 3 dimensional CT images, which are in x,y,z axes, have been used with applying scaling, rotating and translating consecutively.

- Scaling

The scaling of the 3 dimensional object is moving operation of the 1 pixel value from P(x,y,z) coordinates value position in Equation (1) to the S conversion matrix with some defined coefficients to the P' position. With scaling operation, pixel coordinates are approaching or moving away from each other. By this manner, it is aimed to enlarge or reduce the image in x,y,z directions consecutively in Sx, Sy, Sz ratios.

According to the formula stated in Equation (2), the new coordinates of the pixels can be calculated as in Equation (3).

$$S=\begin{bmatrix} Sx & 0 & 0 & 0 \\ 0 & Sy & 0 & 0 \\ 0 & 0 & Sz & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

$$P'=SP \quad (2)$$

$$P_x' = Sx\,Px_1 \qquad P_y' = Sy\,Py_1 \qquad P_z' = Sz\,Pz_1 \tag{3}$$

- Rotation

The value of 3 dimensional objects' in P(x,y,z) position and pixel value in the rotation of the axis just like in Equation (4) is by determining the new coordinates and obtaining P'(x_1, y_1, z_1) position. The images'with the R matrix and θ angle in Equation (5) and the rotation around z axis is ending up with P=>P' conversion. For rotation around x and y axes, assigning of x_1=x ve y_1=y is carried out and for the other axes, the same equalities have been used.

$$x_1 = x\cos\theta - y\sin\theta \qquad y_1 = x\sin\theta - y\cos\theta \qquad z_1 = z \tag{4}$$

$$R=\begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5}$$

- Translation

Translation is the operation of moving from P(x, y, z) position to P'(x_1, y_1, z_1) position without changing the pixel values without changing their values. The translation operation is being carried out to Equation (7) with the T conversion matrix stated in Equation (6). As stated in Equation (8), the pixel value in Px_1, Py_1, Pz_1 position is being shifted in an amount of t_x, t_y, t_z and is being moved to P_x', P_y', P_z' which is the new position.

$$T=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ t_x & t_y & t_z & 1 \end{bmatrix} \tag{6}$$

$$P' = TP \tag{7}$$

$$P_x' = t_x + Px_1 \qquad P_y' = t_y + Py_1 \qquad P_z' = t_z + Pz_1 \tag{8}$$

In this study, the radiologist manually segmented the CT images that belongs to different people.

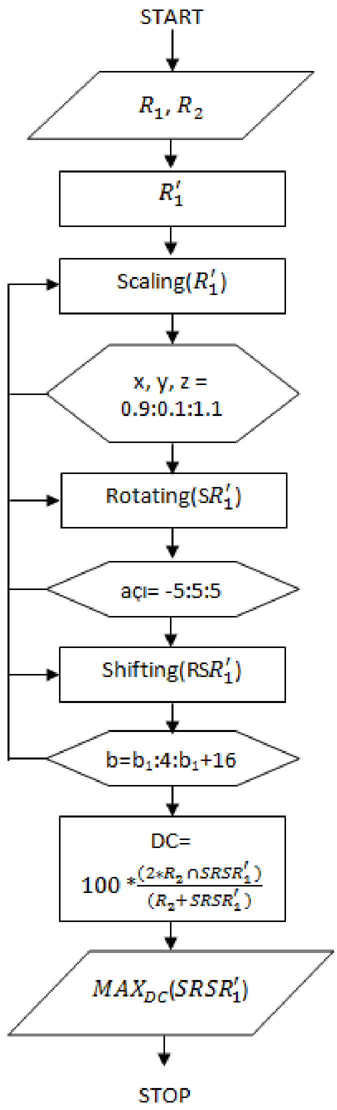


Figure 1. The algorithm steps of the liver images to calculate the overlapping amount

The P_1 and P_2 edges have been manually drawn by two different peoples' CT images by the experts. Connected Componenet Analysis has been applied to P_1 and P_2 images and labelled images have been obtained. P_1' subimage has been obtained from original P_1 image by determining the border point of the Liver images. Magnification or reducing operations have been applied on x,y,z axes lying between 0.9 and 1.1, a step size of 0.1 to the P_1' image. The obtained scaled SP_1' images' similarity has been explored with the P_2 image. RSP_1' is obtained by rotating in (+) and (-) directions the scaled SP_1' image is in the interval of -5 and 5 with a change of 5°. The similarity of the final image with P_2 has been explored. RSP_1' image where a rotating operation has been applied, with a step size of 4 pixels, 16 pixels have been shifted separately on x, y, z axes. The starting point for shifting operation has been determined for the border pixels' starting point at P_2 image. DC (dice coefficient) has been calculated as similarity index in every step. The $TRSP_1'$ image which has the highest DC value has been obtained. Dice coefficient is calculated as follows:

$$DC=\frac{\text{number of true positives}}{\text{number of positives} + \text{number of false positives}} \tag{9}$$

In Equation (9), number of positives is the total number of the image pixels having intensity of 1. Number of true positives is the total number of first and second images' of pixels that have intensity of 1. It uses AND operator on first and second images. Number of false positives is the number of image pixels that have intensity of 1 in second the image but zero in first the image. Its traditional formula is shown in Equation (10). I1 and I2 are the two CT images which their similarities are compared.

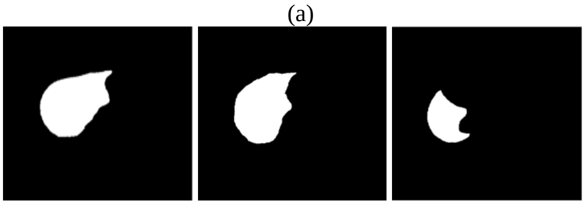
$$DC=\frac{|I_1 \cap I_2|}{|I_1|+|I_2 \setminus I_1|} \tag{10}$$

3. Conclusions

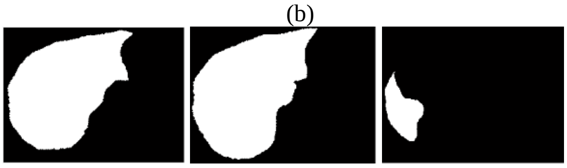
In this study, by the aid of geometrical transformation methods, the similarity of the liver images has been explored.

Belonging to two different patients' CT images', three different slices have been obtained and have been presented in Figure 2(a) and 2(b). After the scaling, rotation and shifting operations that have been carried out with the subimages that have been obtained, three different slices presented by the overlapping of the two images have been presented in Figure 2(c).

(a)



(b)



(c)

Figure 2. (a) R_1 original image, **(b)** R_2 original image, **(c)** the overlapping image after the geometric conversion has been applied.

In order to make a numerical evaluation Dice coefficients have been calculated and shown in the second column of Table 1. These values are obtained by calculating the average of DC values for an each Case. Cases at the first column of Table 1 show the CT images numbers of the patients. The angle values for the rotation geometrical operation are shown in the third column of Table 1. Scaling coefficients in x, y and z axes are shown in the fourth, fifth and sixth column of Table 1, respectively.

Table1. Dice Coefficient, Rotation Angle, Scaling Coefficients in x-axis, y-axis and z-axis

	DC	RA	x-Scale	y-Scale	z-Scale
Case 1	0.7040	0 -5 0 0 -5 5 5 0 0 0	1.0 1.1 1.1 1.1 1.1 1.1 1.1 1.1 1.1 1.1	1.0 1.1 1.1 0.9 1.1 1.1 1.1 1.0 1.1 1.0	1.0 1.1 1.1 1.1 1.1 1.1 1.1 1.1 1.1 0.9
Case 2	0.6520	-5 0 0 -5 0 0 0 -5 0 -5	1.1 1.0 1.1 1.1 0.9 0.9 0.9 1.1 1.1 1.1	1.1 1.0 1.0 1.1 0.9 1.1 1.1 1.1 0.9 1.1	1.1 1.0 0.9 1.0 0.9 1.1 1.1 0.9 1.1 1.1
Case 3	0.7430	0 0 0 0 -5 0 0 0 0 0	1.1 1.1 1.0 0.9 1.1 1.1 1.1 1.0 1.0 0.9	1.1 1.0 1.0 0.9 1.1 1.0 1.0 0.9 1.1 0.9	1.1 0.9 1.0 0.9 1.0 1.1 1.1 0.9 0.9 1.0
Case 4	0.7070	0 -5 0 0 5 0 5 0 5 0	1.1 1.1 0.9 1.0 1.1 1.1 1.0 1.1 1.1 1.0	0.9 1.1 0.9 1.0 1.1 1.1 1.1 1.1 1.1 1.0	1.1 1.0 0.9 1.0 1.1 1.1 1.1 1.0 0.9 0.9
Case 5	0.7160	-5 0 -5 5 0 0 0 0 0 0	1.1 0.9 1.1 1.1 1.0 0.9 0.9 0.9 0.9 0.9	1.1 0.9 1.1 1.1 1.0 0.9 0.9 0.9 0.9 0.9	1.1 0.9 1.0 1.1 1.0 1.1 1.1 0.9 0.9 0.9
Case 6	0.7250	5 0 0 0 0 0 0 0 0 0	1.1 0.9 1.1 1.1 0.9 1.0 1.0 1.1 1.1 1.1	1.1 1.1 1.0 1.1 0.9 1.0 1.0 1.1 0.9 1.1	1.1 1.1 1.1 1.1 1.1 1.0 0.9 1.1 1.1 1.1
Case 7	0.7360	5 0 0 5 0 0 0 0 0 0	1.1 0.9 1.1 1.0 0.9 1.0 1.0 0.9 0.9 0.9	1.1 1.1 1.0 1.1 0.9 1.0 1.0 0.9 1.1 0.9	1.1 1.1 1.1 1.1 1.1 0.9 1.0 0.9 0.9 0.9
Case 8	0.7450	0 -5 0 0 0 0 0 0 0 0	1.1 1.1 1.0 1.1 0.9 1.1 0.9 1.0 1.0 1.1	1.0 1.1 0.9 1.1 0.9 1.1 0.9 1.0 1.1 0.9	1.1 0.9 0.9 1.0 0.9 1.1 0.9 1.0 0.9 0.9
Case 9	0.7420	0 0 0 5 0 0 0 0 0 0	1.1 1.1 1.0 1.1 0.9 1.1 0.9 1.0 1.0 1.0	1.1 0.9 1.1 1.1 0.9 0.9 1.1 1.1 1.0 0.9	1.1 1.1 0.9 0.9 0.9 1.1 0.9 0.9 1.0 0.9
Case 10	0.6010	0 -5 0 0 0 0 0 0 0 0	1.1 1.1 0.9 1.0 0.9 1.1 0.9 1.1 1.0 1.0	1.0 1.1 0.9 1.0 0.9 1.1 0.9 0.9 0.9 1.0	0.9 1.1 1.0 0.9 0.9 1.1 0.9 0.9 0.9 1.0

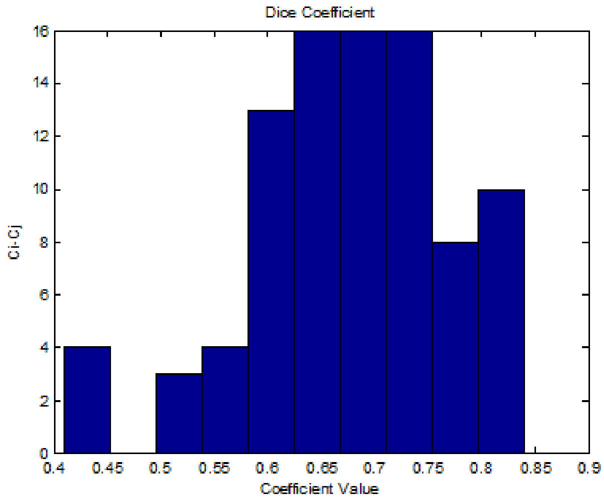


Figure 3. Dice coefficient histogram

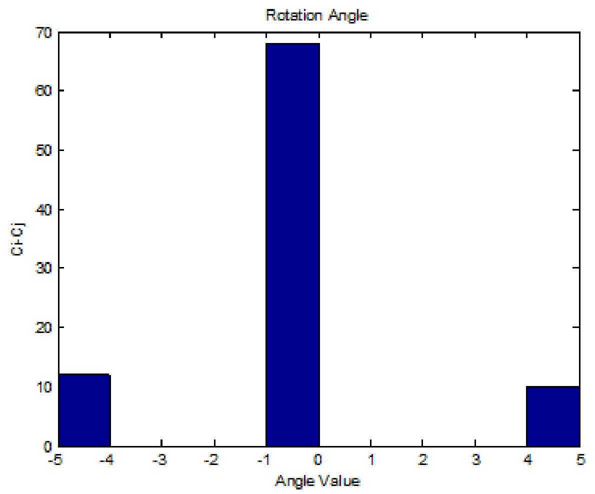
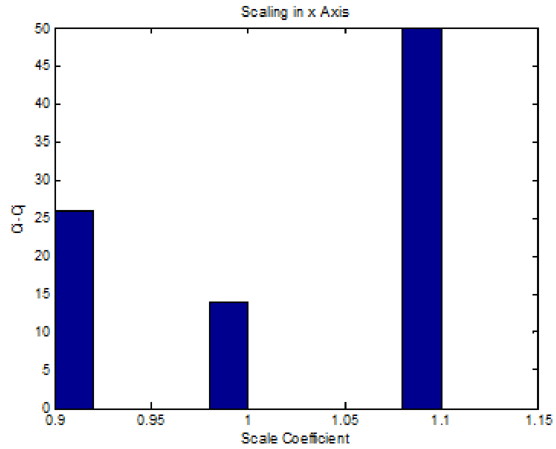
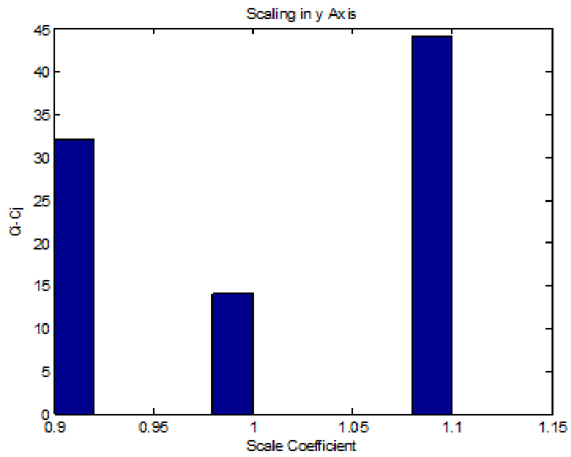


Figure 4. Rotation angle histogram

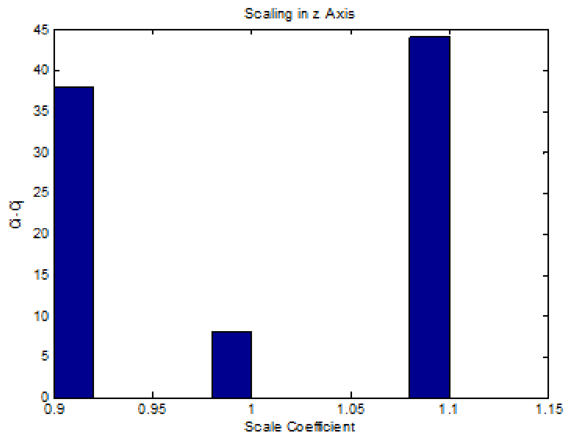
The histogram related to Dice coefficient has been shown in Figure 3. According to the coefficient scattering graph, the liver images that belong to different people (Ci-Cj) show a similarity of 67∓0.09%. Figure 4 is showing the scattering of the angle values for the rotation geometric operation. In Figures 5(a), Figure 5(b) and Figure 5(c) consecutively, the operations that have been carried out on x, y, z axes have been presented



(a)



(b)



(c)

Figure 5. Scaling coefficients histogram **(a)** in x axis, **(b)** in y axis, **(c)** in z axis

4. References

[1] Linguraru, M.G., Sandberg J.K., Li Z., Shah F., Summers R.M., "Automated Segmentation and Quantification of Liver and Spleen from CT Images using Normalized Probabilistic Atlases and Enhancement Estimation", Medical Pyysics, 37(2):771-783, 2010.

[2] Campadelli P., Casiraghi E., Pratissoli S., "A Segmentation Framework for Abdominal Organs from CT Scans", Artificial Intelligence in Medicine, 50:3–11, 2010.9.

[3] Li C., Wang X., Li J., Eberl S., Fulham M., Yin Y., Feng D.D., "Joint Probabilistic Model of Shape and Intensity for Multiple Abdominal Organ Segmentation from Volumetric CT Images", IEEE Journal of Bıomedical and Health Informatics, 17(1): 92-102, 2013.

[4] Chen X., Udupa J.K., Bagci U, Zhuge Y., Yao J., "Medical Image Segmentation by Combining Graph Cuts and Oriented Active Appearance Models", IEEE Transactions on Image Processing, 21(4): 2035-2046, 2012.

[5] Wolz R., Chu C., Misawa K., Fujiwara M., Mori K., Rueckert D., "Automated Abdominal Multi-Organ Segmentation with Subject-Specific Atlas Generation", IEEE Transactions on Medical Imaging, 32(9): 1723-1730, 2013.

[6] Linguraru M.G., Pura J.A, Chowdhury A.S., "Summers R.M., Multi-Organ Segmentation from Multi-Phase Abdominal CT via 4D Graphs using Enhancement, Shape and Location Optimization", Medical Image Computing and Computer-Assisted Intervention – MICCAI , 13(Pt 3): 89–96, 2010.

[7] Palabaş T., Osman, O., Ergin T, Teomete U., Dandin Ö., "Automated Segmentation of the Injured Liver", *Medical Technologies National Conference (TIPTEKNO)*,2015, DOI: 10.1109/TIPTEKNO.2015.7374590.

[8] Dandin, Ö., Teomete, U., Osman, O., Tulum, G., Ergin, T., Sabuncuoğlu, M.Z., "Automated Segmentation of the Injured Spleen", International Journal of Computer Assisted Radiology and Surgery, 2015, DOI:10.1007/s11548-015-1288-9.

[9] Selver, M. A., Kocaoğlu, A., Doğan, H., Demir, G. K., Dicle, O. ve Güzeliş, C., "Nakil Öcesi Verici Değerlendirmeleri için Otomatik Karaciğer Bölütleme Yöntemi", Hastane ve Yaşam, Ocak 2008, 80-87.

[10] Çınar, A. ve Durkaya, A. K., "Karaciğer Görüntüsü Ayıklamak için Genişleme Prensipli Yaklaşım", *National Conference on Electrical Electronics and Computer Engineering-ELECO 2010*, December 2010, 588-591.

A Real-Time Infant Health Monitoring System for Hard of Hearing Parents by using Android-based Mobil Devices

Faruk AKTAS[1], Emre KAVUS[1], Yunus KAVUS[1]

[1]Kocaeli University, Department of Biomedical Engineering, Kocaeli, Turkey
faruk.aktas@kocaeli.edu.tr, emrekavus@gmail.com, yunuskavus@gmail.com

Abstract: *In this study, a real-time infant monitoring system by using android-based mobile devices is developed and implemented in order to be used especially for hard of hearing parents. An Arduino Leonardo board has been used in the system design along with body temperature sensor, sound detection sensor, finger heartbeat detector, and humidity sensor. In order to notify alarm conditions to the parents, an android-based application has been developed. It is observed that the data collected from the sensors are monitored in real-time and the alarms are set off successfully when abnormal conditions occurred.*

Keywords: *Infant monitoring systems, hard of hearing parents, microcontroller and android-based application.*

1. Introduction

Recently, with the development of technology, home healthcare and remote monitoring of physiological data have gained importance. It is a popular implementation to track home healthcare of patients, particularly babies.

An infant monitoring system basically includes sensors and a microcontroller. A biomedical sensor device is capable of sensing several vital physiological and physical data (ECG, body temperature, SpO_2, heartbeat, blood pressure, wetness etc.) from human bodies or environment and sending them to the microcontroller by using analog or digital outputs. It is important that the monitoring system generates alarms in abnormal conditions. Various wireless infant monitoring systems were previously proposed for different purposes [1-13].

One of the most important physiological data to track in such monitoring systems is body temperature, i.e. infant fever. Changes in body temperature of babies have a key role in the diagnosis and treatment of diseases. In particular, rapid febrility in babies can cause vital damage. Therefore, the body temperature should be continuously monitored. The inability to adjust the temperature of the environment may cause excessive perspiration or cooling for premature or weak-born babies. The maximum body temperature range should be 36-38 ^{o}C for these babies. Another crucial parameter to track is heartbeat rhythms. Cardiac arrhythmia can cause sudden deaths of infants [4], so continuously monitoring of the infant's heartbeat rhythm may be required. A finger heartbeat sensor is a low-cost, noninvasive and user-friendly device for monitoring heartbeat rhythms. It is also important to monitor bedwetting and perspiration of babies. When the parents are too late to intervene to bedwetted babies, intertrigo problems may emerge. In this case, the baby may be unrestful, sleeping disorders and febrility may occur. Excessive perspiration may cause the infant be dehydrated, resulting in illness or exacerbation of existing disease. In all of those cases, the sound level of baby crying is a kind of natural alarm. In particular, the detection of a baby crying is therefore very important especially for hard of hearing parents.

The major advantage of homecare systems for infant monitoring is that these systems can automatically collect physiological data without the requirement for parents to constantly check infants, and can generate an alarm for abnormal conditions. It is easy for healthy parents to react immediately to those alarm conditions. However, it may not be possible for hard of hearing parents to react a baby crying or a voice alarm instantaneously. Therefore, alarm notifications must be visual or vibrant for those parents.

In this study, a real-time infant monitoring system by using a microcontroller and android-based mobile devices is developed and implemented in order to be used especially for hard of hearing parents. With the developed system, the physiological signals collected from an infant body are continuously monitored, and an alarm is generated in abnormal conditions. An Arduino Leonardo board has been used in the system design along with a body temperature sensor, a sound detection sensor, a finger heartbeat detector, and a humidity sensor. In order to notify alarm conditions to those parents, an android-based application which is executable on all Android-based smartphones has been developed. Notification of alarm

Received on: 05.02.2017
Accepted on: 14.03.2017

situations has been successfully provided via a vibrating smartwatch, SMS, and LEDs (Light Emitting Diode) using Arduino board and android-based applications. The experimental studies show that the developed system provides a time-saving implementation for home care infant monitoring systems.

The paper is organized as follows: In section 2, components and architecture of the developed system are explained, and the design and implementation of the system are presented in section 3. The paper is concluded in section 4 with final remarks.

2. System Architecture

The real-time infant health monitoring system for hard of hearing parents has been designed and implemented by using Arduino Leonardo boards and android-based mobile devices. The developed system includes sensors, an android-based smartphone, a vibrating Smartwatch for those parents, and a microcontroller that evaluates the received data from the sensors and generates an alarm when emergency conditions occur. It is important to note that an the android-based application namely “Infant Monitoring” has been developed in order to send alarm events to Smartwatch. The architecture of the system is illustrated in Figure 1.

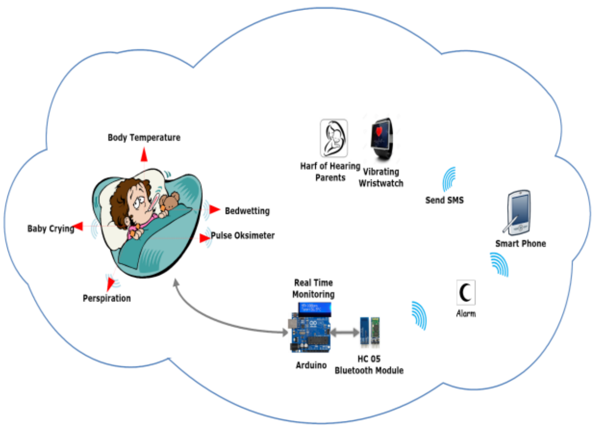


Figure 1. Infant monitoring system for hard of hearing parents

In the following, components of the developed system are explained in detail.

2.1. Components of the System

The system consists of several components;

- Atmega32u4 microcontroller on Arduino Leonardo board
- LCD screen
- Android smartphone
- Vibrating smartwatch
- HC-05 Bluetooth module
- Body temperature sensor
- Finger heartbeat sensor
- Humidity sensor

The Arduino Leonardo development board have been used to collect data from the sensors and evaluates these data to create cases which trigger the alarm. It is a microcontroller board based on the Atmega32u4. It has 20 digital input/output pins (of which 7 can be used as PWM outputs and 12 as analog inputs) which enable to collect and process many physiological data [14].

In order to monitoring body temperature of the infant, a waterproof version of the DS18B20 temperature sensor have been used [15]. The analog body temperature raw data collected from the sensor is converted into the actual temperature data using a microcontroller.

The KY-039 finger heartbeat sensor compatible with Arduino boards has been used to monitor the infant heartbeat. The sensor consists of two components, namely an infrared phototransistor (sensor) and an infrared LED (IRLED). The heartbeat is obtained according to the amount of light passing through the finger between the IRLED and the sensor [16].

The T1592 humidity sensor has been used for wetness sensing. This sensor can be used to monitor both infant bedwetting and perspiration.

The KY-037 sound detection sensor has been used in order to detect baby crying. Sound detection sensor has two outputs: Analog output (A0) is real-time microphone output port. D0 port generates a high and low-level signal when the sound volume reaches the threshold. The threshold-sensibility can be adjusted via the potentiometer on the sensor.

Bluetooth is a wireless standard (IEEE standardized Bluetooth as IEEE 802.15.1) for exchanging data over short distances (using short-wavelength UHF radio waves in the ISM band from 2.4 to 2.485 GHz) from fixed and mobile devices. The HC-05 Bluetooth module has been used for the notification of alarm situations to send wireless to the parents via the android-based smartphone [17].

Smartwatch has been used to inform parents with a vibration alarm. This device is able to communicate with the android-based smartphone via Bluetooth technology. Before using this device, “Bluetooth notice” (BTNotification) application is required to be installed on an android-based smartphone in order to automatically synchronize the device. Smartwatch has a user interface to display “message” and “remote notice” from a smartphone. Additionally, it is able to receive notification of each application or send a reminder, including SMS alerts, and other messages/reminders. The Smartwatch can be charged via micro USB and can standby time up to 120 hours. This low-cost device with vibration features is an ideal choice in the notification of the parents.

3. Design and Implementation

We now present experimental studies the developed system by using aforementioned components and the android-based application. Experimental studies of infant health monitoring system for hard of hearing parents have been carried out in a laboratory environment and physiological data was collected from an adult.

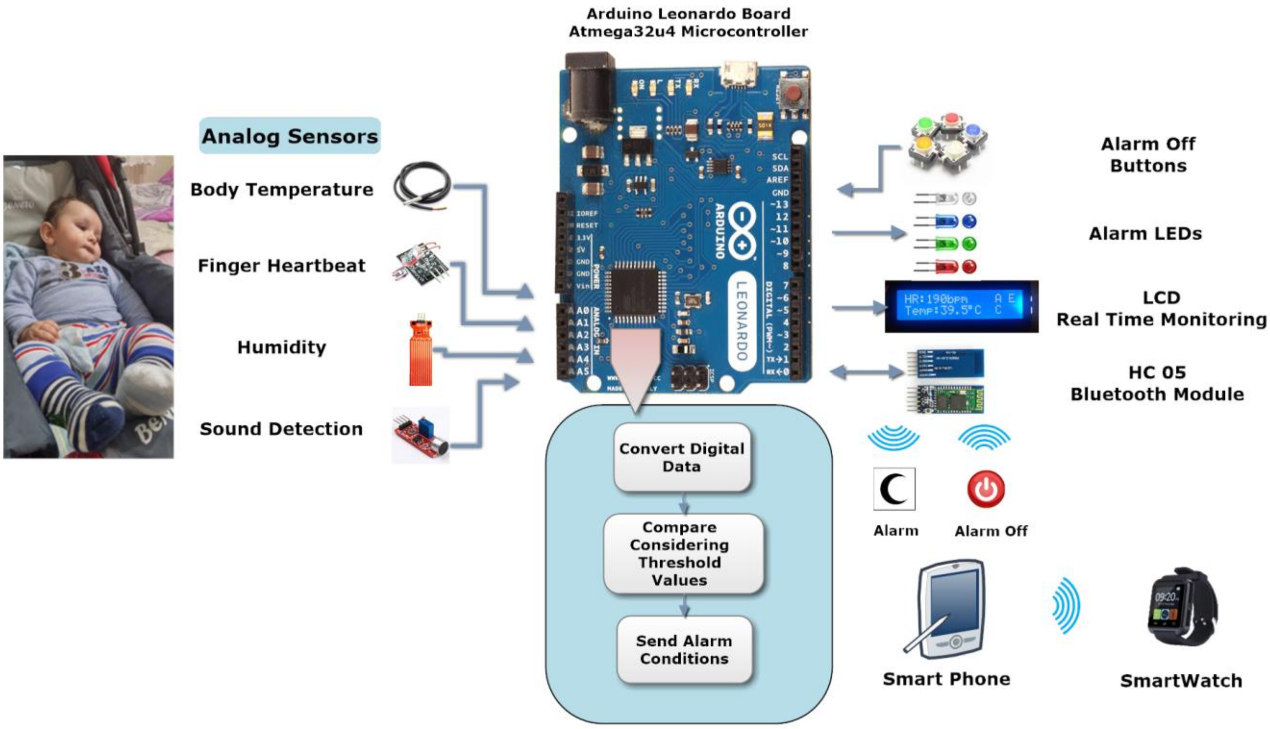


Figure 2. Hardware connection diagram

3.1. Hardware Design

The microcontroller performs three important functions; converting the physiological data collected from the sensors to digital raw data, generating an alarm using digital raw data according to the determined conditions, and providing notification to the parents. In this context, the hardware connection diagram and microcontroller process procedure of the developed system are depicted in Figure 2.

As shown in Figure 2, the sensors (i.e. body temperature, finger heartbeat, humidity, sound detection sensors) are connected to the Arduino Leonardo board via analog inputs. Body temperature sensor is attached either to the armpit or finger of an infant for accurate measurement results. The humidity sensor can be used for two purposes; detecting bedwetting or perspiration. It is placed in the infant's nappy for bedwetting detection, and on the back of infant for perspiration detection. Heartbeat data is collected by attaching the finger heartbeat sensor to the infant's finger. Sound detection sensor is placed at a distance to the infant for baby crying detection.

In the microcontroller firstly the physiological data collected from the sensors are converted to digital raw data using analog-digital converter (ADC) module. ADC module resolution is 10 bits in the Atmega32u4 microcontroller. The raw data collected from the sensors displayed on the LCD screen are normalized considering real values. For example, while the infant body temperature is 36 degrees, the raw data values obtained from the sensor is 75. By the normalization process, the raw data value of 75 is converted to the real value of 36 degrees.

The process after conversion of the analog data to normalized digital data is the determination of the alarm conditions. Special alarm conditions have been determined according to the data collected from sensors in order to alert the family in abnormal situations. The digital data is compared with the corresponding threshold values for alarm conditions. A character code has been assigned for each alarm condition. The codes, cases, and conditions of the alarms are given in Table 1.

Table 1. Alarm codes and alarm conditions

	Alarm code	Alarm Causes	Alarm Condition
1	A	High heartbeat	>135 bpm
2	B	Low heart beat	<80 bpm
3	C	High body temperature	>38 ºC
4	D	Low body temperature	<34 ºC
5	E	Bedwetting or perspiration	Wetness
6	F	Baby crying	Threshold value exceeded 15 times

It can be seen from Table 1 that alarm conditions are generated if the infant's fever reaches 38 ºC (alarm code "C") or falls below 34 ºC (alarm code "D"). Alarm conditions for the data obtained from the finger heartbeat sensor have been determined for high heartbeat (alarm code "A") or low heartbeat (alarm code "B") in accordance with the age range of infants. For example, the 0-5 month baby's heartbeat is 100-160 bpm (bit per minutes), 6-12 months baby's heartbeat is 80-140 bpm [2]. In this study, the boundary values are chosen 80-135 bpm (1-3 years age) bpm for alarm conditions. The alarm thresholds have been determined under different conditions in the detection of bedwetting and perspiration. An alarm is generated when the wetness level reaches different threshold values for bedwetting or perspiration detection (alarm code “E”). For baby crying detection, a sound level threshold is prescribed. If this threshold is exceeded continuously (e.g. more than fifteen times) in a specified time duration (e.g. 10 sec), an alarm is generated (alarm code “F”). This mechanism avoids any false alarms in case of short duration and/or low-level sound changes.

The last process performed by the microcontroller is to notify the alerting cases to the parent. In the event of an alarm, both the related LED blinks and the related alarm code is displayed on the LCD screen. At the same time, alarm information is sent to the android-based smartphone via Bluetooth module to notify the vibrating Smartwatch. The parents can manually turn the alarm off by pressing a related button. Figure 2 shows the LCD screen for a high heartbeat, high body temperature, and bedwetting or perspiration alarm conditions.

In the next section, the developed android-based application for notification of alarm cases to those parents is introduced and explained in detail.

3.2. Android-based Application

We now introduce the android-based application, an important component of the developed system. Android is a mobile operating system developed by Google, based on the Linux kernel and designed mainly for touchscreen mobile devices such as smartphones and tablets. In this study, a user interface on the android-based application has been developed that performs several functions to notify the parents of alarm conditions. The android-based application was developed in the App Inventor web application. The prime advantage of the developed application is that it can work on all Android phones and not only just on a special phone. App Inventor for Android is an open-source web application originally provided by Google and now provided by the Massachusetts Institute of Technology (MIT) [18]. The android application is connected to Arduino Leonardo board via Bluetooth module. When it receives alarm information from the Arduino, the android application sends an SMS containing alarm information to the parents. At the same time, the phone vibrates and gives an alarm message for a second. The application user interface is shown in Figure 3.

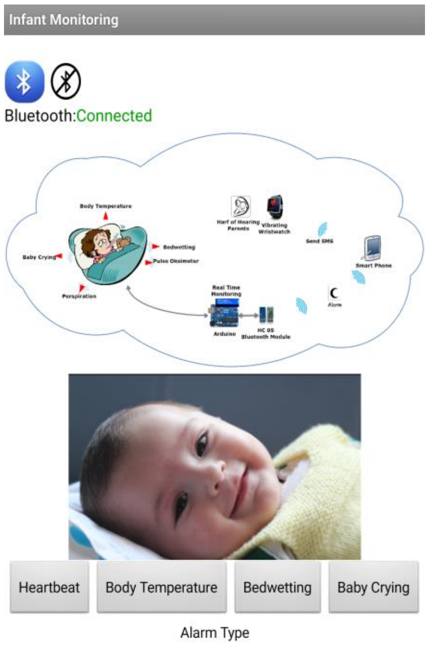


Figure 3. Android application user interface

As shown in Figure 3, there are two buttons on the user interface of application related to the Bluetooth connection. Bluetooth connection defaults to “Not Connected”. When the Bluetooth button (blue icon) is clicked, active Bluetooth devices in the vicinity of the smartphone are listed. After the HC-05 Bluetooth module is selected, Arduino Leonardo, is connected to the smartphone via Bluetooth and "Connected" is displayed (green text color) on the application screen. If the disconnect button (black-white icon) is clicked, the Bluetooth is disconnected and "Not Connected" is displayed (red text color) again on the application screen. With “BTNotification” application, it is possible to synchronize smartphone with smartwatch via Bluetooth. So when there is a notification on the smartphone, the smartwatch vibrates.

In the event of any alarm, the microcontroller system sends an alarm code to the android-based application via Bluetooth. When this code is received, the alarm notification procedure is activated in the application. A “Message Dialog Box” notification is generated according to the incoming alarm code and the smartphone vibrates for one second. The corresponding alarm information is displayed in the message box. At the same time, an SMS (containing the alarm condition) is sent to the previously defined (parents’) phone numbers. The SMS provides a vibration by creating a notification on the smartwatch synchronized to the smartphone via Bluetooth. The android-based application and smartwatch screen for baby crying and bedwetting alarm is shown in Figure 4 and Figure 5, respectively.

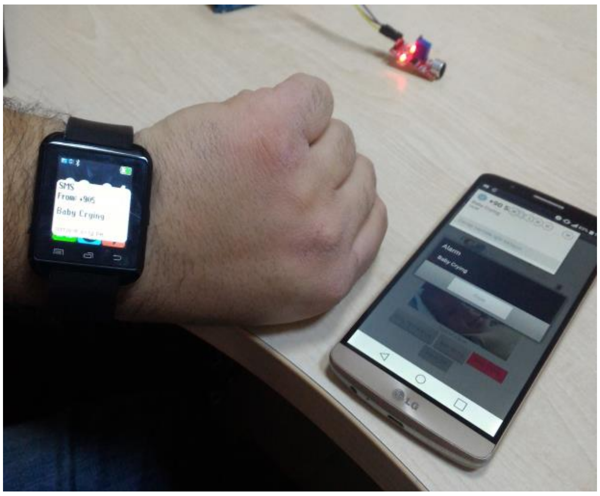


Figure 4. Smartphone and smartwatch screen for baby crying alarm

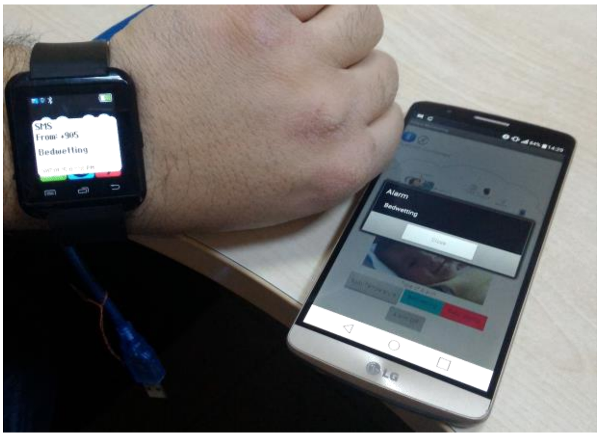


Figure 5. Smartphone and smartwatch screen for bedwetting alarm

As shown in Figure 4 and Figure 5, in the case of an alarm, the related button background color changes in the "Alarm Type" section, and the previously invisible "Alarm Off" button becomes visible. When the "Alarm Off" button is clicked, information related to closing the alarm is sent to the microcontroller via Bluetooth and the alarm is deactivated and the related led is turned off. In the application, the related button returns to the previous state and the "Alarm stop" button becomes invisible again. The aim of the notifications is to inform the parents in case of any alarm. According to the alarm, the parents can either treat the baby themselves or take it to a medical center.

4. Conclusions

In this paper, we have presented a real-time infant monitoring system for hard of hearing parents, consisting of sensors (finger heartbeat, body temperature, humidity and sound detection), a microcontroller and android-based mobile devices (smartphone and smartwatch). In particular, a system has been developed that monitors both physiological data collected from infants and creates alarms for abnormal conditions.The designed and implemented system, developed on the Arduino Leonardo board, has been used in order to data collect data from the sensors and to create alarm cases by evaluating these data. Low-cost vibrating Smartwatch compatible with android-based smartphones has been used for notification of alarms to the parent. From the implementation results, it is observed that the data collected from the sensors are monitored real-time and that the alarms determined when abnormal conditions occur are notified successfully.

The implemented system is designed as an open system for improvement and other desired sensors can be easily added to the system. This study can be used not only for hard of hearing parents but also for healthy parents. In addition, the developed system, which is suitable for monitoring adult patients, can be used for general purposes. We remark that it would be useful to have some usability tests including an average of measurements (e.g. from 10 babies) for the purpose statistical evaluations of the proposed system, which will be considered as a future work.

5. Acknowledge

We thank all the reviewers for their invaluable comments.

6. References

[1] H. Cao, L-C. Hsu, T. Ativanichayaphong, J. Sin, J-C Chiao, "A non-invasive and remote infant monitoring system using CO2 sensors", in *IEEE Sensors 2007 Conference*, Atlanta, USA, 2007, pp. *989-992.*

[2] E. Saadatian, S. P. Iyer, C. Lihui, O. N. N. Fernando, N. Hideaki, A. D. Cheok, A. P. Maduapperuma, G. Ponnampalam, Z. Amin, “Low-Cost Infant Monitoring and Communication System”, in *2011 IEEE Colloquium on Humanities, Science and Engineering Research (CHUSER 2011)*, Penang, CHINA, 2011, pp. *503-508.*

[3] S-H. Cheng, J-C. Huang, C-J Lin, "A real-time Location and Infant Monitoring System Based on Active RFID", in *Proceedings of the 2012 International Conference on Machine Learning and Cybernetics*, Xian, CHINA, 2012, pp. *1844-1849.*

[4] W. Lin, R. Zhang, J. Brittelli, C. Lehmann, " Wireless Infant Monitoring Device for the Prevention of Sudden Infant Death Syndrome", in *11th International Conference & Expo on Emerging Technologies for a Smarter World (CEWIT)*, New York, USA, 2014.

[5] S. Mohajerani, S. A. H. Moosavi, R-A. Rihawi, B. Ahmed, A. N. Bhat, R. Y. Kamal, "A Cloud-Based System for Real-Time, Remote Physiological Monitoring of Infants", in *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Abu Dhabi, UAE, 2015, pp. *565-569.*

[6] H. Zhou, B. Goold, "A Domestic Adaptable Infant Monitoring System Using Wireless Sensor Networks", in *IEEE 34th International Performance Computing and Communications Conference (IPCCC)*, Nanjing, CHINA, 2015.

[7] A. G. Ferreira, D. Fernandes, S. Branco, J. L. Monteiro, J. Cabral, A. P. Catarino, A. M. Rocha, "A Smart

Wearable System for Sudden Infant Death Syndrome Monitoring", in *IEEE International Conference on Industrial Technology (ICIT)*, Taipei, TAIWAN, 2016, pp. *1920-1925.*
[8] P. Kumari, P. Goel, S. R. N. Reddy, "PiCam: IoT based Wireless Alert System for Deaf and Hard of Hearing", in *International Conference on Advanced Computing and Communications*, Coimbatore, INDIA, 2015, pp. *39-44.*
[9] M-J. Wu, S-F. Shieh, Y-L. Liao, Y-C. Chen, "ECG Measurement System Based on Arduino and Android Devices", in *International Symposium on Computer, Consumer and Control*, Xian, CHINA, 2016, pp. 6*90-693.*
[10] H. Deng, S. Chen, "Design and implementation of Android-based health and healthcare system", in *International Wireless Communications and Mobile Computing Conference (IWCMC)*, Paphos, CYPRUS, 2016, pp. *164-169.*
[11] S. Agezo, Y. Zhang, Z. Ye, S. Chopra, S. Vora, T. Kurzweg, "Battery-Free RFID Heart Rate Monitoring System", in *IEEE Wireless Health (WH)*, Bethesda, MD, 2016, pp. *136-142.*
[12] V. Wahane, "An Android Based Wireless ECG Monitoring System for Cardiac Arrhythmia", in *IEEE Healthcare Innovation Point-Of-Care Technologies Conference (HI-POCT)*, Cancun, MEXICO, 2016, pp. *183-187.*
[13] P. S. Chowdary, S. Aruna, "Infant Monitoring System", *International Journal of Computer Science and Communication*, vol. *2*, no. *2*, pp. *501-503*, July-December, 2011.
[14] www.arduino.cc/en/Main/ArduinoBoardLeonardo
[15] http://datasheets.maximintegrated.com/en/ds/DS18B20.pdf
[16] http://www.zoebachman.net/itp/?p=190
[17] https://www.itead.cc/wiki/Serial_Port_Bluetooth_Module_(Master/Slave)_:_HC-05
[18] http://appinventor.mit.edu/explore/

Faruk Aktaş received his B.S. and M.S. degrees in Electronic and Computer Education from Kocaeli University, Turkey, in 2008 and 2012, respectively. He is currently pursuing his Ph.D. degree in Biomedical Engineering at the Kocaeli University. His active researches are wireless body area networks, internet of things, microcontrollers and embedded devices.

Emre Kavuş currently pursuing his B.S. degree in Biomedical Engineering at the Kocaeli University. His interests include medical technology, biosensors, medical imaging and wireless networks.

Yunus Kavuş currently pursuing his B.S. degree in Biomedical Engineering at the Kocaeli University. His interests include medical technologies and device design, biosensors and medical imagin

The Role of Sports Participation in Hemispheric Dominance: Assessment by Electrodermal Activity Signals

Serhat ALADAĞ[1], Ayşegül GÜVEN[1], Nazan DOLU[2], Hatice ÖZBEK[2]

[1]Department of Biomedical Engineering, Erciyes University, Kayseri, Turkey
serhat.aladag@titck.gov.tr, aguven@erciyes.edu.tr
[2]Department of Physiology, Medical School, Erciyes University, Kayseri, Turkey
dolu@erciyes.edu.tr, haticeozbek380@gmail.com

Abstract: *This study aims at evaluating hemispheric differences between active sportsmen (n=17) and sedentary control subjects (n=21) using Electrodermal Activity (EDA), which is a physiologic measure of emotional sweating. Following the denoising of the acquired EDA records, feature extraction functions were utilized for each record to generate a feature vector, containing 14 parameters per record. Statistical significances of the differences and similarities between feature vectors were determined by unpaired t-test ($p<0.05$ and $p>0.95$). Findings from this study show that the agreement of hemispheric dominance test with hand preference is significantly lower for active sportsmen in comparison to control group. This implies that participation in sport activity may play a facilitative role in a more equally weighted development of both hemispheres by somewhat decreasing the hemispheric lateralization.*
Keywords: *signal processing; feature extraction; statistical comparison; denoising methods, electro dermal activity.*

1. Introduction

Electrodermal Activity (EDA) is defined by the electrical activity that is stemming from the electrochemical interactions between sweat glands and, epidermal and dermal layers of the skin.

Electrodes positioning on the certain locations of the skin surface can measure this activity [1,2].

EDA is a physiologic measure for the assessment of the sympathetic activity. Therefore, this technique has found widespread applications in the clinical neurophysiology and psychophysiology.

In most cases, EDA signals are not used solely. Instead, they are commonly utilized as a supporting parameter for polygraph tests. Nevertheless, EDA records have been evaluated in the investigation of diseases such as hemispheric asymmetry, anxiety, depression and schizophrenia [3]. For example, Bakker et al. (2011) have classified parameters according to their effects on the dynamic stress levels in daily life by applying median filtering and symbolic assembly methods [4].

Several other studies have focused on combined assessment of physiological signals together with EDA [5]. Furthermore, EDA have been taken into account along with electrocardiogram (ECG) and electromyogram (EMG) records for the real-time stress level monitoring during driving [6].

Note that EDA terms a general notion to describe all electrical activity occurring on the skin level. These activities comprise active and passive electrical properties of the skin and its secondary structures.

There are two common techniques in the evaluation of skins electrical properties, known as endosomatic and exosomatic methods. The exosomatic method can either use a direct current (DC) or alternating current (AC) through a circuit consisting of a galvanometer, electric battery and human body to measure changes in EDA; however DC currents are used more often than AC currents. In contrast the second method, the endosomatic method, uses a human body and galvanometer circuit to measure changes in resting electromotive force, or voltage of the skin [7-13].

In the present study, we hypothesized that sport participation reduces hemispheric lateralization. The aim of the study was to test this hypothesis by employing exosomatic technique.

2. Methods

2.1. Subjects and Data Recording

In this study, data recording and experimental procedures were approved by the Ethics Committee of Clinical Sciences at Erciyes Univesity. All data recordings were carried out at Brain Dynamics Laboratory, Physiology Department, Medical School, Erciyes University.

Received on: 24.01.2017
Accepted on: 13.03.2017

Among the 38 healthy individuals (age: 20±0.4, mean±SD) who agreed to participate, 17 were sportsmen (8 right-handed and 9 left-handed) and the remaining were sedentary control subjects (11 right- handed and 10 left- handed).

Sportsmen's expertisement branches which are especially needs team coordination just as football (n=6), basketball (n=4) and volley (n=7).

EDA records were acquired using MP30 (Biopac Systems Inc., USA) via electrodes placed on the palmar regions of the distal phalanxes of the thumb and index fingers of both hands, with a sampling frequency of 200HZ is seen Fig. 1.

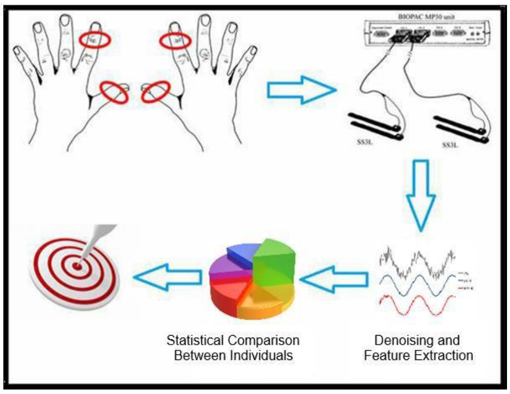


Figure 1. Schematic flow-chart of the data acquisition and processing steps. Electrode positions are indicated by circles on the fingers.

Signal processing, denoising and feature extraction operations were carried out in Erciyes University Biomedical Engineering Department.

2.2. Phasic and Tonic EDA Records

In the present study, resting state (tonic) EDA signals were recorded during 120 seconds, whereas phasic EDA signals were recorded during the Raven Progressive Matrices (RSPM) protocol, which requires subjects to answer 20 questions. Prior to signal recording, hemispheric dominance and hand preference tests were applied per subject.

In addition, SPO2 and ECG wristband (W/Me, Phyode Inc., USA) was monitored to ensure that subjects were not under the influence of hormonal, emotional, physiological and environmental effects. This wristband is responsive against the physiological changes occurring in response to subjects' mood. If any signs of internal or external excitatory factors were detected, subjects were given enough time before appropriate conditions for data recording were reached.

2.3. Determination of Hemispheric Dominance and Hand Preference

Human brain is "cross-wired" so that the left hemisphere controls the right handed side of the body and the right hemisphere controls the left handed side of the body.

Hand preference has been commonly taken as a basis for the determination of hemispheric dominance for many years. Recent reports also provide support for the relationship between handedness and hemispheric dominance.

In the present study, cerebral dominance was determined according to the subjects' score on the Annett Hand Preference Questionnaire and Hemispheric Dominance Test. [14]. Results from these tests are further compared with the RSPM procedure which is a measure of general cognitive abilities and can also be used in the determination of hemispheric dominance. Findings from these test were compared with the verbal declaration of hand preference made by participants.

2.3.1. Annett Hand Preference Questionnaire

Hand preference is related with hemispheric dominance. Therefore, we utilized Annett Hand Preference Test to compare our results.

This test directs 13 questions to determine handedness of an individual. Assessment is based on the choice of hand preference for different daily activities such as brushing teeth, using spoon and other handtools, opening jar, striking a match etc.

If score falls between 13-17 points, subject is considered as right handed. Whereas, if it is between 18-32 points, individual is ambidextral. Aby scores above 32 indicates left handedness [14].

2.3.2. Hemispheric Dominance Test

Although Annet's test is a commonly preferred method for hand preference determination, this evaluation is based solely on the script. To strengthen this evaluation, we also applied RSPM, in which the hand preference is determined using visual test items.

2.3.3. Raven's Progressive Matrices

Raven's Progressive Matrices method dates back to past in the assessment of general cognitive abilities. However, it can also be employed to determine hemispheric dominance.

This test includes 20 visual test items. For each test item, the subject is asked to identify the missing element that completes a certain pattern. Fig.2 shows a sample question from RSPM test.

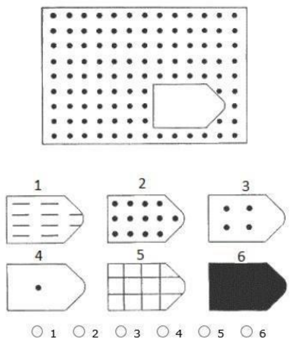


Figure 2. A sample question from RSPM procedure

2.4. Signal Denoising

Biopotential signals attain low frequency components. For example, EDA signal has a characteristic signal frequency within a range of 0.0167 to 0.25 Hz. Therefore, it is highly affected by line frequency noise interference [15].

Based on the comparison of signal denoising performance measures by Aladag et al. (2015), Singular Spectrum Analysis (SSA) was qualified as a substantially efficient method for the noise removal of EDA signals thanks to its high decomposition sensitivity. Therefore, this technique was utilized to eliminate power-line interference from the noisy-raw EDA signals are seen above section of Fig. 3.

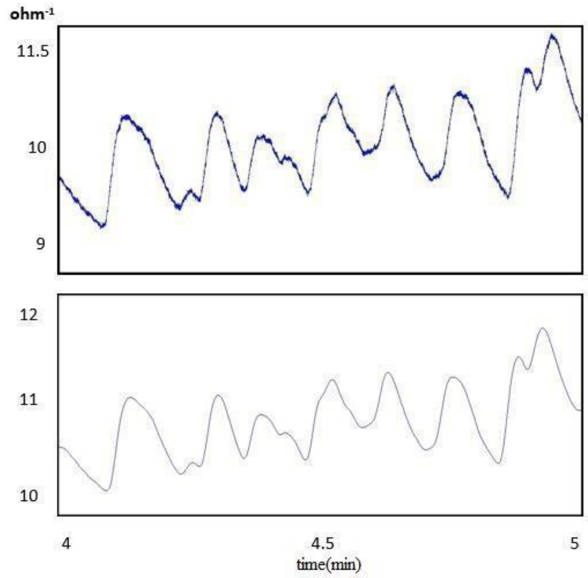


Figure 3. Original EDA signal from a representative subject (upper panel) and the same signal after denoising with SSA (lower panel).

2.5. Feature Extraction

Parameter identification is the first step to create a feature vector that enables between-groups comparisons for detecting significant differences. Essential statistical measures such as average, median, standard deviation, variance, maximum, minimum, range, mode values, entrophy, skewness, kurtosis, peaks of the EDA signal can be incorporated into feature vector as parameters.

In the present study, measures for uncertainty such as entropy, skewness and kurtosis are also included in the feature vector along with essential statistical measures. Furthermore, mean signal power and root mean square (RMS) are added into feature vector as additional signal characterization metrics. Feature vector generation method applied in this study has been described in detail elsewhere [16]. And feature extraction function was applied phasic and tonic records, after the signal denoising with SSA.

3. Determination of Feature Vector Identification Way

Parameter identification is an essential step to determine statistical differences or similarities between individuals. Therefore, we have chosen parameters with utmost attention to provide a sufficient separation power for statistical significance.

In the present study, three different methods were employed on the denoised signals to determine statistical differences ($p<0.05$) and similarities ($p>0.95$). These techniques are detailed within this section.

3.1. Application of SSA Followed by Feature Extraction Function

The Singular Spectrum Analysis (SSA) technique is a powerful technique of time series analysis, incorporating the elements of classical time series analysis, multivariate statistics, multivariate geometry, dynamical systems and signal processing [17].

The aim of SSA is to decompose original time series into the sum of a small number of independent and interpretable components such as slowly varying trends, oscillatory components and random noise. Decomposition of signal into its principal values is carried out with respect to the selected window length (Fig4).

$$\begin{cases} \text{Stage 1 : Decomposition} & \begin{cases} \text{Step 1 : Embedding} \\ \text{Step 2 : Singular Value Decomposition (SVD)} \end{cases} \\ \text{Stage 2 : Reconstruction} & \begin{cases} \text{Step 1 : Grouping} \\ \text{Step 2 : Diagonal Averaging} \end{cases} \end{cases}$$

Figure 4. A short description of the SSA technique (for more information see [18])

After several trials, a window length of 7 was considered suitable to obtain an acceptable signal. Following this, subsignals were acquired, which were denoted as "R" in Fig. 5. Feature extraction function was applied to R2+R3 as seen Fig 6, R4+R7 and other combinations are also applied.

A "p value" lower than 0,05 indicates statistical significance, whereas "p value" higher than 0,95 implies similarity. Application of SSA did not reveal any statistical significance or similarity.

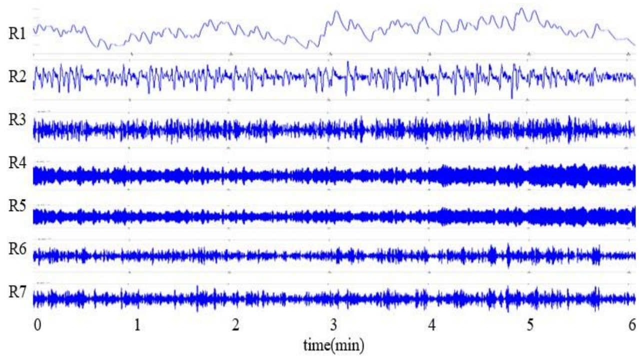


Figure 5. SSA decomposition with a window length of 7 and demonstration of components.

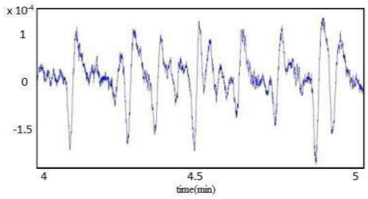


Figure 6. SSA applied for the R2 an R3 summation

3.2. Standalone Employment of Feature Extraction

Using SSA method only, feature extraction function was applied on denoised signal. However, no statistical meaning was observed among the participants with this approach.

3.3. Application of DWT Followed by Feature Extraction Function

Discrete wavelet transforms (DWTs), analyze signals and images into finer octave bands progressively. This multiresolution analysis enables the detection of the patterns that are not visible in the raw data. This method can also be used to reconstruct signals (1–D) and image (2–D) approximations that contains desired features only, and to compare the distribution of energy in signals across frequency bands [19]. In the present study, we applied 10th Degree Haar Filtering DWT to the denoised signals (Fig. 7). This transform is calculated using the equation 1.

When DWT is applied to a certain sampled function s(t), this function becomes decomposed as the addition of a set of signals, namely wavelet signals: An approximation of a signal at a certain decomposition level n (a_n) plus n detail signals (d_j with j varying from 1 to n). The mathematical expression characterizing this process is given by equation 1, where $(\alpha_i)^n$, $(\beta_i)^j$, are the scaling and wavelet coefficients, $\varphi^n(t)$, $\psi^j(t)$ are the scaling function at level n and wavelet function at level j, respectively, and n is the decomposition level [20-22].

$$s(t) = \sum_i \alpha_i^n \cdot \varphi_i^n(t) + \sum_{j=1}^{n} \sum_i \beta_i^j \cdot \psi_i^j(t) = a_n + d_n + \ldots + d_1 \qquad (1)$$

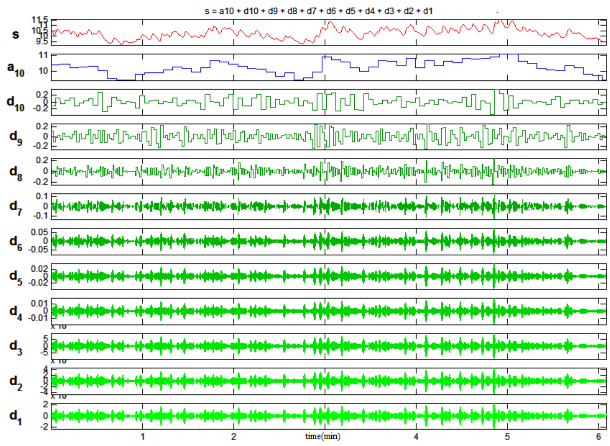


Figure 7. 10th degree Haar Filtering DWT

Next, we applied feature extraction function to reconstructed signal from 10th degree DWT's subsignals (Fig. 8).

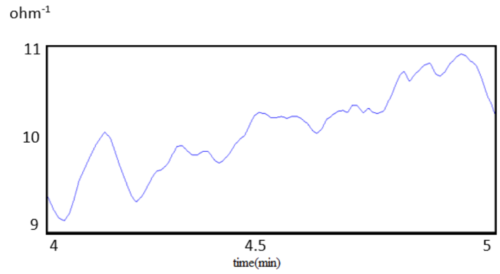


Figure 8. Reconstructed signal from 10th degree DWT's subsignals (a10+d9+d8+d1)

We found statistical differences and similarities using this method. Findings from this method explained in the results section.

4. Results

4.1. Annett Hand Preference Questionnaire Result

Test results from the control group revealed that among the sedentary participants who declared right-handedness, 18,2% were left handed. Ratio of the subjects whose hand preference was rejected by the tests was 10% for left handedness in the sedentary group.

In sportsmen the ratio of the rejection were higher. Among these participants who declared right-handedness only 37,5% was scored as right handed in tests, whereas 62,5% was characterized as ambidextral. Similarly, ratio of the subjects who are characterized as left handed and those as ambidextral were 33,3% and 66,7% respectively.

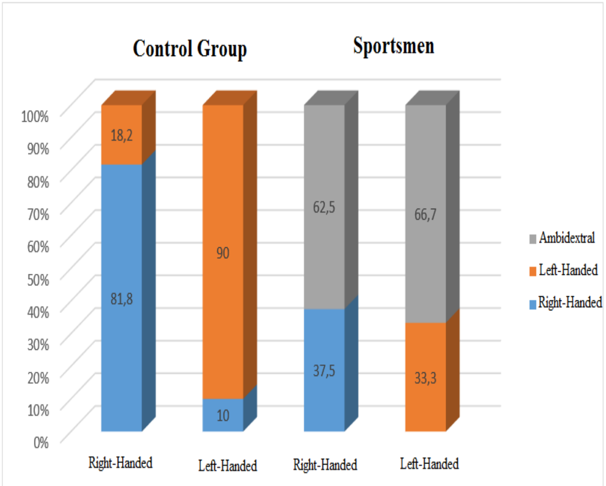


Figure 9. Annett hand preference questionnaire results

4.2. Hemispheric Dominance Test Result

Hemispheric dominance test results showed that 80% of right-handed individuals use left hemisphere, whereas 90,9% of left-handed individuals use right hemisphere (Fig. 10).

In sportsmen, 21,4% of right-handed individuals appeared to use their left hemisphere, whereas 7,1% them use right hemisphere and 71,5% use both hemispheres. On the other hand, 14,2% of left-handed individuals appeared to use right hemisphere, whereas 7,1% use left hemisphere and 78,7% use both hemispheres (Fig. 9).

Note that such less emphasized hemispheric lateralization of sportsmen constitutes a support for the hypothesis. However, a statistical confirmation is required for its acceptance.

Due to the involvement of stimulation, EDA values from phasic records were higher than those obtained from tonic records.

Statistical significance of phasic records was equal to that of tonic records.

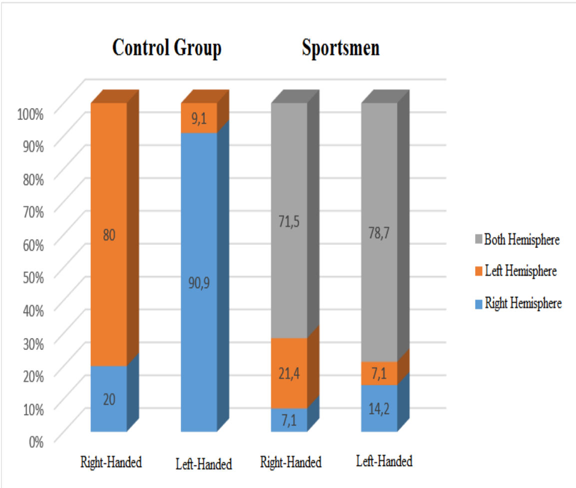


Figure 10. Hemispheric dominance test results

5. Conclusions

Effect of sports participation on hemispheric dominance was evaluated in the present study.

EDA records were denoised using signal processing methods to extract reliable information. Next feature extraction functions were applied on denoised EDA signal for to create feature vectors for comparison. 14 parameters were incorporated into this vector with regard to their statistical significance.

In accordance with the previous findings from literature [2], our findings show that hemispheric difference is more emphasized in sedentary subjects in comparison to sportsmen.

Handedness indicated by Annett and hemispheric dominance tests were in good compliance with the hand preference of the sedentary subjects in their daily lives. However, such compliance was less prominent for sportsmen. This implies that sportsmen follow a more balanced strategy for hemispheric recruitment.

The evaluation of this compliance was based on the statistical significances ($p<0.05$) of the differences or similarities ($p>0.95$) by applying unpaired t-test between feature vectors those are obtained using EDA signals.

According to Table 1, sedentaries left hand entropy is different than their right hand entropy. However, for sportmen, left hand entropy equaled to the right hand entropy, indicating a difference between sedentary subjects and sportsmen.

According to Table 2, right hand significance value of left handed sedentary subjects is different from that of left handed sportmen’s for Kurtosis, indicating a difference between sedentary subjects and sportsmen.

Significant differences between sportsmen and sedentary subjects were indicated by entrophy, skewness and kurtosis calculations in Table 3 for those subject shown in Table 2.

Findings from this study accept our hypothesis, suggesting that the agreement of hemispheric dominance test with hand preference is significantly lower for active sportsmen in comparison to control group.

This implies that participation in sport activity may play a facilitative role in a more equally weighted development of both hemispheres by decreasing the hemispheric lateralization.

Table 1. Comparison of intra-group significance values

Feature	Right Hand and Left Hand Comparison of Right Handed Sedentery	Right Hand and Left Hand Comparison of Left Handed Sedentery	Right Hand and Left Hand Comparison of Right Handed Sportmen	Right Hand and Left Hand Comparison of Left Handed Sportmen
Entrophy	0,83	0,69	1,00	1,00

Table 2. Comparison of sportmen and sedentary subjects with identical hand preferences.

Feature	Right Hand Comparison of Right Handed Sedentery and Sportmen	Left Hand Comparison of Right Handed Sedentery and Sportmen	Right Hand Comparison of Left Handed Sedentery and Sportmen	Left Hand Comparison of Left Handed Sedentery and Sportmen
Kurtosis	0,38	0,23	0,05	0,13

Table 3. Comparison of sportmen and sedentary subjects with difference hand preferences.

Feature	Left Hand Comparison of Right Handed and Left Handed Sportmen	Right Hand Comparison of Right Handed and Left Handed Sportmen	Left Hand Comparison of Right Handed and Left Handed Sedentery	Right Hand Comparison of Right Handed and Left Handed Sedentery
Entrophy	0,70	0,70	0,05	0,05
Skewness	0,04	0,04	0,87	0,16
Kurtosis	0,23	0,28	0,05	0,05

6. Acknowledgements

This work was supported by The Erciyes University Scientific Research Projects Department under grant FYL-2015-5684.

The authors are thankful to Agah KARAKUZU for his technical contributions.

7. References

[1] Bouscein W. Electrodermal Activity. Newyork, Plenum Pres. 1992, 1-372.

[2] Dolu N, Özbek H, Determination of the Hemispheric Difference at Attention Level in Sportsman, Erciyes University, Kayseri, 2009

[3] Dolu N, Yüksek A, Sizer A, Alay M. Arousal and continuous attention during Ramadan intermittent fasting. J Basic Clin Physiol Pharmacol. 2007;18(4):315-22.
[4] Bakker J., sensor data, ICDMW '11 Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops.
[5] Begum S., Ahmed MU., Funk P., Xiong N., and Schéele B.V, (2009), Case-Based Decision Support System for Individual Stress Diagnosis Using Fuzzy Similarity Matching, The Journal of Computational Intelligence (CI), vol 25, nr 3, p180-195.
[6] Rigas G., Goletsis Y., and Fotiadis DI., Real-Time Driver's Stress Event Detection, IEEE Transactions on Intelligent Transportation Systems 2012; 13:221-234.
[7] Stefan Schmidt and Harald Walach, (2000), Electro Dermal Activity (EDA) Research – State of the art measurement and techniques for Parapsychological purposes.
[8] Boucsein W. Electrodermal Activity. New York: Plenum Press. 1992.
[9] www.stanfordchildrens.org, "What is a superficial SDYstdegree burn?", Standford's Children Health Lucile Packard Children's Hospital, Standford
[10] Ethan Leng, Mihir Mongia, Charles Park, Tiffany Varughese, Andrew Wu, SMART Belt: A Low-cost Seizure Detection Device, Rice University
[11] Geddes LA, Baker LE. Principles of Applied Biomedical Instrumentation, 3rd ed. New York: John Wiley and Sons.1989.
[12] Qiao ZG, Morkrid L, Grimnes S. Three-electrode Method to Study Event-related Responses in Skin Electrical Potential, Admittance and Blood Flow. Med Biol Eng Comput 1987;25: 567-72.
[13] Stern RM, Ray WJ, Davis CM. Psychophysiological Recording. New York: Oxford: Oxford Universty Press. 1980: 196-210.
[14] Uslu, R. (1996). Neurological Soft Signs In Children With Psychiatric Disorders. Child Psychiatry Thesis. Department of Child Psychiatry, Ankara University.
[15] Aladağ S., Güven A., Özbek H., Dolu N., "A Comparison of Denoising Methods for Electrodermal Activity Signals", Tıp Teknolojileri Kongresi 2015, MUĞLA, TURKEY, 15-18 Ekim 2015, ss.213-216
[16] Aladağ S., Güven A., Dolu N., Özbek H., (2016). "Measuring Electrodermal Activity To Determining Sympathetic Activity In Sportsman And Feature Extraction With Signal Processing Methods", Tıp Teknolojileri Kongresi, TIPTEKNO, ANTALYA, TURKEY, 27-29 Ekim 2016, ss.98-101
[17] Hassani H., "Singular Spectrum Analysis: Methodology and Comparison", Journal of Data Science 5(2007), 239-257
[18] Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. (2001). Analysis of Time Series Structure: SSA and related techniques, Chapman & Hall/CRC, New York - London.
[19] Chaovalit P., Gangopadhyay A., "Discrete Wavelet Transform-Based Time Series Analysis and Mining",ACM Computing Surveys, Vol. 43, No. 2, Article 6, Publication date: January 2011.
[20] Burrus, C.S.; Gopinath, R.A.; Guo, H.: "Introduction to Wavelets and Wavelets Transforms. A Primer." Prentice Hall, 1997.
[21] Mallat, S.: "A Wavelet Tour of Signal Processing, Third Edition". Academic Press, dec. 2008.
[22] Chui, C.K.: "Wavelets: A Mathematical Tool for Signal Analysis", SIAM, 1997.

Serhat ALADAĞ worked in the field of Clinical Engineering at Dr. Abdurrahman Yurtaslan Oncology Training and Research Hospital, Ankara, Turkey for a short period. In the same field, he worked in Turkey Public Hospitals Agency (TPHA), Ankara, Turkey for the following 2.5 years. Later on, he changed his field of study and started working at Turkish Medicines and Medical Devices Agency (TMMDA), Ankara, Turkey in May 2016. He is currently working for TMMDA at Quality Management and R&D Laboratory Department. He received his BSc and MSc degrees in Biomedical Engineering from Erciyes University, Kayseri, Turkey, in 2013 and 2017, respectively.

Ayşegül GÜVEN received respectively her BSc, MSc and PhD degrees from Erciyes University Electrics and Electronics Engineering in 1996, 1999 and 2006, respectively. She is currently an associate professor at the Biomedical Engineering Department of Erciyes University. Her current research interests include signal processing, medical imaging and classification with artificial neural networks. She authored or co-authored in 60 papers published in scientific journals and conference proceedings combined.

Nazan Dolu worked at the Erciyes University from 1993 to 2016. Then she retired as Prof. of Department of Psychology, Erciyes University, Kayseri in 2016. She received her medical training from Cerrahpaşa Faculty of Medicine, Istanbul University, İstanbul in 1989 and completed her post-doctorate research in Psychology at the Erciyes University, Faculty of Medicine, Kayseri in 1996. Her specific teaching and research interests are neuroscience, electrophysiology, neurophysiology and molecular biology. She authored or co-authored in 144 papers published in scientific journals and conference proceedings combined.

Hatice ÖZBEK graduated from Biology Programme, Faculty of Science and Letters at Erciyes University, Kayseri, Turkey in 1999. She worked as teacher at Mürşitali Primary School, Iğdır, Turkey. Then she received her MSc degree from Erciyes University, Kayseri in 2008. She is currently working as a teacher at Sümer Secondary School, Kayseri, Turkey.

BUILDING NEUROCOMPUTATIONAL MODELS AT DIFFERENT LEVELS FOR BASAL GANGLIA CIRCUIT

Rahmi Elibol[1], Neslihan Serap ŞENGÖR[1]

[1]Istanbul Technical University, Electronics and Communication Engineering, Istanbul, Turkiye
rahmielibol@itu.edu.tr, sengorn@itu.edu.tr

Abstract: *The target of computational neuroscience studies can be considered in two-folds: understanding the connection between the physiology and functional aspects of the brain to develop new approaches for diagnosing and treatment of neurological disorders and behavioral deficits and understanding mind and consciousness to develop new intelligent technologies. The methods and approaches used in computational neuroscience have to overcome the complexity of the system in all aspects. So, different methods and approaches are developed for different scales not only for observing the phenomena, but also for modeling. In this paper, an approach is proposed to build a connection between different levels of modeling. A simple, linear system will be shown to give an understanding of the working principle of basal ganglia circuit which is modeled with a detailed spiking neural network approach. First, spiking neural network of basal ganglia circuit will be introduced and the role of dopamine on its functioning will be shown; then a simple linear system model will be given, and the relation between two models will be explained. The aim of this work is to show that even a simple model which is not sufficient for detailed understanding of the neuronal process, could give a coarse understanding of a complex phenomenon. Such simple models could be used as a starting point in building complex models and also can be benefited for implementing intelligent technologies.*
Keywords: *Basal ganglia circuit, spiking neural network, mass model, dopamine.*

1. Introduction

In computational neuroscience literature, there are numerous models of neurons and neural structures at different levels. One reason of this diversity is the collection of data at different levels using different measurement tools and methods. While based on single neuron measurements, it is possible to obtain data to build a detailed model of a neuronal behavior based on the role of ion channels [1,2], it is also possible to pinpoint the regions of the brain that are active during a task by fMRI (functional magnetic resonance imaging) and obtain neural field model which can mimic the collective activity of neurons at a specific region during a specific task [3]. The resolution of measurements also depend on the scale. While, temporal resolution of single neuron measurements and EEG (electroencephalogram) / LFP (local field potential) are better, spatial resolution of fMRI is superior to other techniques. Thus the models corresponding to different levels have to cope with all these different scales and the dynamics of the brain is either modeled by a set of nonlinear, ordinary differential equations or partial differential equations. Besides these, there are hybrid models, where some structures are modeled at neuron level; others are modeled as mass model [4,5].

Of course the role of different aspects on neurological disorders and diseases is another reason of this diversity. While mutual activity of neurons is responsible for some processes and the malfunctioning in their collective behavior give rise to deficits, the activity of ion channels and the concentration of ions and neurotransmitters are important in other cases. So models differ as they target all these different phenomena at different levels. This variety of models is needed since all provide information necessary to understand the complexity of cognitive processes and neurological disorders and diseases. In computational neuroscience, the level of the model built has to be decided considering the experimental results to be used and the cognitive process dealt with. In some cases, models at different levels should be considered together to have a better understanding.

There is another aspect which should be noted for the models in computational neuroscience other than building the connection between the physiology and functional aspects of the brain. As pointed out in abstract, computational neuroscience also focuses on understanding mind and consciousness to develop new intelligent technologies. For this aspect, the simplicity of the models is crucial, since implementation on a hardware and real time applications is possible only if the computational burden is manageable. Even though there are some attempts to develop special hardware for neural structures as

Received on: 31.01.2017
Accepted on: 13.03.2017

SpiNNaker [6], neuromorphic hardware as Neurogrid [7], the scale of model should be kept small to use the models in mobile and robotics applications [4]. Thus, while modeling the behavior of a group of spiking neurons, the number of neurons in the model is not same as the number of neurons in the neuronal structure to be modeled, but scaled to a number that is capable to display the behavior. Also, the membrane potential of a neuron is modeled by either by first order dynamical systems as integrate and fire models [8,9] or second order dynamical systems as Izhikevich model [10], even though a detailed model could be obtained by adding ion channel dynamics to Hodgkin-Huxley model [11]. Thus, in neurorobotics applications and in developing new learning rules simple models are preferred, rather than detailed models.

In this paper, the objective is whether it is possible to foresee the behavior of a complicated computational model by a simple one. If this is fulfilled than it would be possible to build a connection between different levels of modeling and a tool can be developed to ease detailed modeling. Since models at each level are versatile as they point different aspects of the neural phenomenon, the aim is not to replace a detailed model by a simple one, but to use a coarse approach to understand a complex but detailed one. To show the possibility of such an approach basal ganglia network will be considered, and it will be modeled by spiking neural models and by a simple linear system.

Modeling basal ganglia network has been considered in computational neuroscience literature extensively [9,11-16] due to its role in voluntarily action selection, reward related learning and in neurological deficits and diseases as Parkinson's disease, Huntington's disease, and in behavioral deficits as addiction. Especially, models of basal ganglia network are developed to understand deep brain stimulation [2]. In recent years more attention is paid to the role of basal ganglia circuits in high level cognitive processes as decision making [17], substance dependence [18,19].

In the following section, a brief introduction to the basal ganglia circuits will be given, especially focusing on the role of dopamine in action selection. Then, the proposed spiking neuron model will be introduced and a simple linear mass model will be given. The simulation results obtained using BRIAN simulator and XPPAUT will be given and these results will be discussed. It will be shown that a connection between the the firing rate of the spiking neuron model and dynamic behavior of simple linear model can be drawn.

2. Basal Ganglia Circuit

Basal ganglia circuits proposed to have important role in motor activation and cognitive processes [20] especially their role in reward based learning and decision making pointed in various works [17,21-23]. Impairment of basal ganglia circuits manifest deficits in motor actions observed in neurodegenerative diseases such as Parkinson's and Huntington's disease, and also cause behavioral deficits observed in attention deficit hyperactivity disorder (ADHD), obsessive-compulsive disorder (OCD) and addiction [19,24-27]. These behavioral disorders and motor movement disorders are treated by deep brain stimulation (DBS), a well-known treatment of Parkinson's disease [28,29]. The role of basal ganglia in psychiatric disorders is considered more recently [30,31], and their treatment by DBS makes basal ganglia a target for functional and restorative neurosurgery [32,33].

In computational neuroscience, modeling basal ganglia circuits is an important subject and many models are proposed. While mostly focus more on the role of basal ganglia circuits in voluntary movement and action selection [9,11-16], there are also models for reinforcement learning [34-42]. Some models deal with the malfunctioning of basal ganglia circuits and how treatments can be developed [2,43-46]. Most of these work except [13,42,45] focus on simple spiking neuron networks. Here, both mass model simpler than the ones in [13,42,45] and spiking neuron networks will be considered.

Striatum is considered as the input structure of basal ganglia and it is together with Subthalamic nucleus (STN), Globus pallidus (internal(GPi), external(GPe) and ventral pallidum) and substantia nigra (pars compacta and pars reticulata) [47,48] form the direct, indirect and hyper-direct pathways of cortico- striatal circuit [20,24]. Normally, direct and indirect pathways are at an equilibrium state. Little perturbations on the output of basal ganglia circuit which correspond to the GPi/SNr (substantia nigra pars reticulata), result in the selection of an action. The role of the hyper-direct pathway is to perform the fine tuning between several possible output choices which are conducted by direct or indirect pathway [49]. Dopamine from substantia nigra pars compacta and ventral tegmental area modify the activation in basal ganglia circuit by acting on striatum.

The afferents of striatum are mainly cortical pyramidal neurons located in layer V and occasionally in layer III [50]. During motor activation, posterior putamen and the dorsolateral anterior putamen receive inputs from motor and motor association cortex [21,51].

Striatum is mostly composed of medium spiny neurons (MSNs) which comprise 80-95% [47] (90 - 95 % [48]) of striatum, remainder is mostly interneurons. Even though MSNs are structurally homogeneous they have different chemical properties and are classified according to their response to neurotransmitter dopamine (DA) [48]. The two most effective groups are MSN with D1 type and D2 type receptors. D2 type receptors are more abundant and they are considered to be promoting the selection of the latent reinforcers [52].

The stimulation of the D1 type DA receptors cause neuronal excitation while the stimulation of the D2 type receptors cause neuronal inhibition. Both D1 and D2 type receptors exist together on the membrane of a neuron and their net effect drives the neuronal output. While D1 type MSNs inhibit GPi neurons and form direct pathway, D2 type MSNs inhibit GPe neurons and form indirect pathway. The direct pathway promotes the action initiation and selection whereas indirect pathway prevents actions [21,53]. The role of DA on striatum behaviour is vital, many neurological diseases and disorders are due to malfunctioning of dopamine neurons in striatum [2,16,46].

The well-known and extensively studied basal ganglia action selection circuit [24] has two main pathways: direct and indirect. Both pathways start with the stimulus from cortex to the neostriatum (caudate and putamen - Str) and unite again at the output nucleus of the basal ganglia, GPi/SNr.

The direct pathway is responsible for action selection while the indirect pathway is responsible for the inhibition of the unwanted actions. Direct pathway starts with the glutamatergic projections from the cortex to the MSNs of striatum. Some of the striatal neurons have direct GABAergic (gamma aminobutyric acid) projections to the GPi. The connection between GPi and thalamus (THL) is inhibitory while the connection from thalamus to cortex is excitatory. The net result of the direct pathway is the inhibition of the inhibition on thalamus so the motor cortical areas are stimulated which is called disinhibition in the neuroscience literature [47,48]. In the indirect pathway different type of striatal neurons have inhibitory GABAergic connections to GPe. GPe inhibits STN while STN stimulates GPi via glutamatergic connection. The net result of the indirect pathway is the disinhibition on the GPi so thalamus is inhibited and motor cortical areas are less stimulated.

In addition to the direct and indirect pathways, there is another excitatory connection between the cortex and the STN and this is called the hyper-direct pathway [49]. Via hyper-direct pathway cortical activity is transferred to GPi over STN and the striatal pathways are short-circuited. By the help of the excitation from the STN, already tonically active GPi inhibits thalamus stronger. For action selection, ventral oralis nuclei of thalamus is active [54] and its efferent is layer IV of supplementary motor area of cortex [21,48,54,55].

3. Computational Models of Basal Ganglia Circuit at Different Levels

In this section, two different levels of modeling basal ganglia circuit will be considered and a spiking neuron model based on Izhikevich type neurons will be introduced first. Then, a simple continuous time differential equation set will be used to define a mass model of basal ganglia circuit. Such mass models for basal ganglia circuit have been given previously [13,42,56,57], but all these were composed of discrete time, nonlinear dynamical systems. These two model focus on the effect of dopamine on action selection and the role of dopamine is modeled by a parameter.

3.1. Spiking neural network model

In this work, a computational model of basal ganglia-thalamocortical circuit for action selection which is shown in **Figure 1** is first built using point neurons. The point neuron model used in forming these groups is Izhikevich neuron [58] and the equations governing the neuron model with the reset condition are given in **Equation 1** and **Equation 2**, respectively. While forming the model the neural activation pattern of each

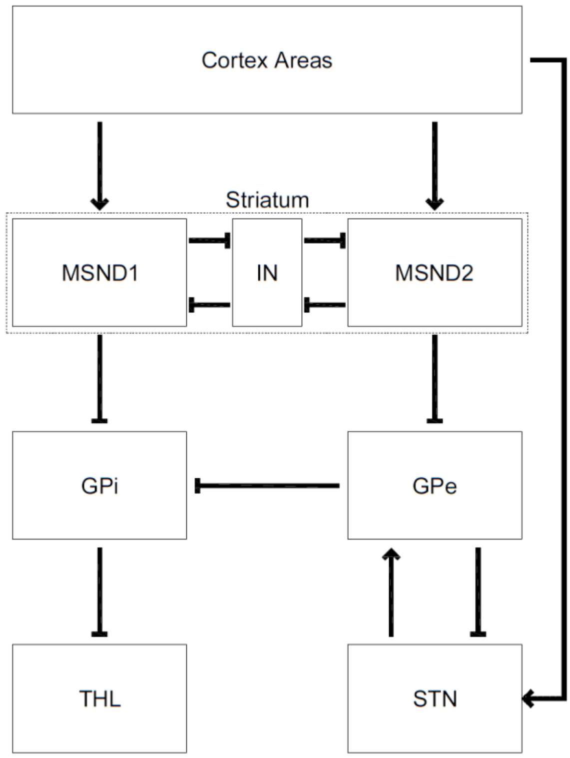


Figure 1. Cortex - Basal Ganglia and Thalamus circuit.

component of basal ganglia circuit is considered and to model different patterns different parameters of Izhikevich neuron model are used. These parameter values are given in the upper part of **Table 1**.

$$v' = 0.04v^2 + 5v + 140 - u + ge - gi \qquad (1)$$
$$u' = a(bv - u)$$

$$\text{If } v > 30 \text{ mV, then } \begin{cases} v \leftarrow c \\ u \leftarrow u + d \end{cases} \qquad (2)$$

In **Equation 1**, g_e and g_i represents the excitatory and inhibitory connections to the neuron. As a neuron is connected to numerous neurons sum of excitatory and inhibitory neurons effects the behavior of the neuron [59]. The dynamics of synapses and the parameter values are given in **Equation 3**, **Equation 4** and **Table 2**, respectively.

$$g'_x = -\frac{g_x}{\tau_{syn}}, x \in \{e, i\} \qquad (3)$$

Here index e corresponds to excitatory, and i corresponds to inhibitory connections.

$$v^{(j)} > V_{thr} \text{ then } g_x^{(k)} \leftarrow g_x^{(k)} + w_{j-k,x}$$
(4)

As the connection increases when a presynaptic neuron fires, $w_{j-k,x}$ denotes this increase in the weight from neuron j to neuron k either excitatory or inhibitory.

While cortex in the model has 200 excitatory pyramidal neurons which have regular spiking activity. Striatum model consists of three groups of neuron populations: D1 type medium spiny neurons (MSND1), D2 type medium spiny neurons (MSND2) and interneurons (IN).

Table 1. Izhikevich model parameters and connection weights. $\neg$ is used for inhibitory and $\rightarrow$ for excitatory connections.

Parameters	RS	FS
a	0.02/ms	0.1/ms
b	0.25/ms	0.2/ms
c	-65mV	-65mV
d	8mV/ms	2mV/ms
Connections	$w_{j-k,x}$	Probabilities
CRTX $\rightarrow$ MSND1	0.75V/s	0.5
IN $\neg$ MSND1	1V/s	0.5
CRTX $\rightarrow$ MSND2	0.75V/s	0.5
IN $\neg$ MSND2	1V/s	0.5
MSND1 $\neg$ IN	1V/s	0.25
MSND2 $\neg$ IN	1V/s	0.25
MSND2 $\neg$ GPe	1V/s	0.25
STN $\rightarrow$ GPe	1V/s	0.25
GPe $\neg$ STN	1V/s	0.25
CRTX $\rightarrow$ STN	1V/s	0.25
MSND1 $\neg$ GPi	0.75V/s	0.25
GPe $\neg$ GPi	0.75V/s	0.25
GPi $\neg$ THL	1V/s	0.25

MSND1 and MSND2 neuron populations are composed of 100 neurons and IN group has 25 neurons. MSND1 and MSND2 are modeled as regular spiking neurons like cortex excitatory neurons. The neurons in GPi, GPe, IN and STN are modeled as fast spiking neurons with initial values v_i=-75, u_i =-16, while cortex inhibitory neurons have as initial condition v_i=-65, u_i=-15. The connection weights and probabilities for the model in **Figure 1** are given in the lower part of **Table 1**. The number of point neurons considered for all structures are given in **Table 3**.

In order to model the role of dopamine on action selection, the synaptic connections have to be changed with dopamine and here, this is accomplished as in [60,61]. Thus, the Equations given in **Equation 3**, **Equation 4** will be modulated with dopamine as in **Equation 5**.

$$g'_x = -\frac{g_x}{\alpha.\tau_{syn}}, x \in \{e, i\} \quad (5)$$

Here, with these equations implemented to the model, the effect of Dopamine can be investigated by changing DA parameter. In order to show different effect of dopamine on MSND1 and MSND2, the parameter α =DA is used for MSND1 group and α =1/DA is used for MSND2 group.

Table 2. Synaptic time constant and frequencies of Poisson groups.

τ_{syn}	CRTX	GPe, THL	STN	GPi
10ms	100Hz	150Hz	200Hz	250Hz

Table 3. Number of neurons in each neural population given in Figure 1.

Neural Population	# of neurons	Behaviour
CRTX	200	RS
MSND1	100	RS
MSND2	100	RS
IN	25	FS
GPi	100	FS
GPe	100	FS
STN	100	FS
THL	100	FS

3.2. Mass model

A mass model equations for basal ganglia circuits are formed by linear differential equations given in **Equation 6**. This model is inspired by the firing rate results of spiking neuron model given in previous section and knowledge of state space behavior of linear dynamical systems.

The behavior of cortex areas $crtx$ are modeled by a Heaviside function. Striatum is represented by two state variables msn_{D1} and msn_{D2}, which have afferent exitatory connection from $crtx$ weighted by dopamine level (DA) denoted by w_{DA1} and w_{DA2}, for msn_{D1} and msn_{D2}, respectively. Each of neural structures other than striatum is modeled by a single state variable, thus gp_e, gp_i, stn and thl are represented by a single dynamical variable and they all have afferent and efferent connections, corresponding to the circuit given in **Figure 1**. The role of dopamine on action selection is investigated by changing dopamine level DA from low to high levels where the values of DA are taken as 0.25, 0.5 and 0.75 for low, normal, and high level respectively [62,63].

$$msn_{D1}' = -msn_{D1} + w_{DA1}.crtx$$
$$msn_{D2}' = -msn_{D2} + w_{DA2}.crtx$$

$$gp_e' = -(gp_e - 0.5) - 0,5.msn_{D2} + 0,5.stn$$
$$stn' = -(stn - 1) - 0,5.gp_e + 0,5.crtx$$
$$gp_i' = -(gp_i - 1) - 0,5.msn_{D1} - 0,5.gp_e$$
(6)
$$thl' = -(thl - 1) - gp_i$$
$$w_{DA1} = DA$$
$$w_{DA2} = 1 - DA$$

4. Results and Conclusion

The simulations are done for spiking neural network model and mass model proposed in Section 3. In both cases the role of dopamine on action selection is investigated similarly. The levels of dopamine are changed through time and its effect on the results are shown with the activity of THL population. This activity which will be called THL activity is depicted with raster plot and firing rate for spiking neuron model and with curves obtained by the dynamical system for mass model.

4.1. Spiking neural network results

The simulations for spiking neural network model are carried in Python based simulation environment BRIAN [59]. During the simulations the level of dopamine is modified by changing parameter α in **Equation 5**. As the effect of dopamine on D1 and D2 type receptors is different, and α =DA for MSND1 group and α =1/DA for MSND2 group, to change the value of α, DA is taken as 1 for normal level, 0.9 for low level and 1.1 for high level. This modification is done through time and the level of dopamine is taken to be at normal level for the first 500ms and then switched to high level till 1500ms and then switched to normal level till 2500ms, to low level till 3500ms and again to normal level till the simulation ended at 4000ms. This switching is done to investigate the effect of dopamine level change on the action selection through time. The simulation results are given for a randomly chosen single neuron from MSND1 and MSND2 population and synaptic activity in **Figure 2**, where the effect of dopamine on single neuron activity can be followed. As it can be followed from **Figure 2**, the neuron chosen from MSND1 population is more active when the level of dopamine is high, and the neuron chosen from MSND2 population is more active when the level of dopamine is low. This change in activity is also projected to synaptic dynamics and synaptic activity in both populations show difference as the membrane potentials.

The outcome of the activity of neuronal populations in basal ganglia circuit determines the activity of THL which can be interpreted as the result of action selection. Thus eventhough the single neuron activity of striatum MSN population reflect the effect of dopamine, the overall result of all this neuronal activity is revealed in THL, so the population activity of MSND1, MSND2 and THL are given in **Figure 3**, with raster plots and firing rates of neurons in MSND1, MSND2 and THL populations.

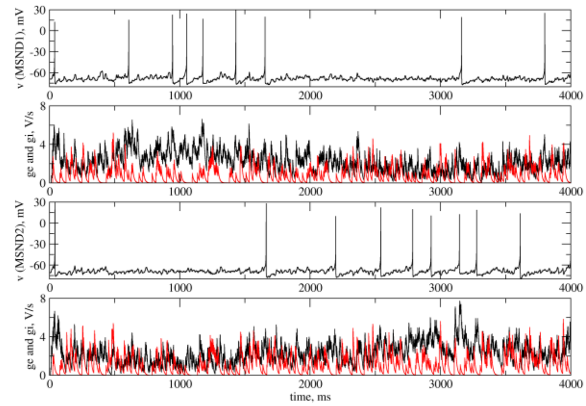


Figure 2. Membrane potential $v(t)$ and synaptic dynamics $g_e(t)$ and $g_i(t)$ for neurons which are randomly selected in MSND1 and MSND2 groups. g_e and g_i is black and red, respectively. DA level is 1 (normal), 1.1 (high), 1 (normal), 0.9 (low) and 1 (normal) for 0-500ms, 500ms-1500ms, 1500ms-2500ms, 2500ms - 3500ms and 3500ms-4000ms time intervals.

In time course of the change of dopamine level can be followed from firing rates easily while raster plot gives the information of neuronal population at neuron level. The firing rates inspired the idea of proposing a simple model for the interaction of neuronal population, where each population activity is denoted by a single dynamical variable as proposed in mass model.

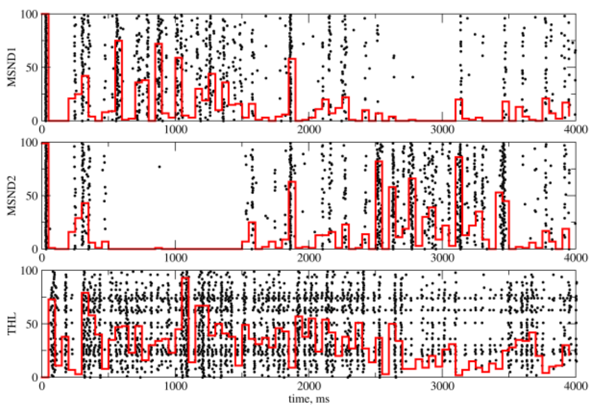


Figure 3. Raster plot and firing rates of MSND1, MSND2 and THL. When DA level is high, THL activity is higher than normal rate and DA level is low, THL activity is lower than normal rate.

4.2. Mass model results

The simulation for the mass model are done using XPPAUT, a tool developed for dynamical system analysis. In order to demonstrate the analogy between the firing rate results of spiking neuronal network and mass model, the time course of dynamical behavior of the variables corresponding to MSND1, MSND2 and THL populations are given in **Figure 4**. Here, again the dopamine level is switched similarly from normal to high, then to normal level followed by low level and finally to normal level by

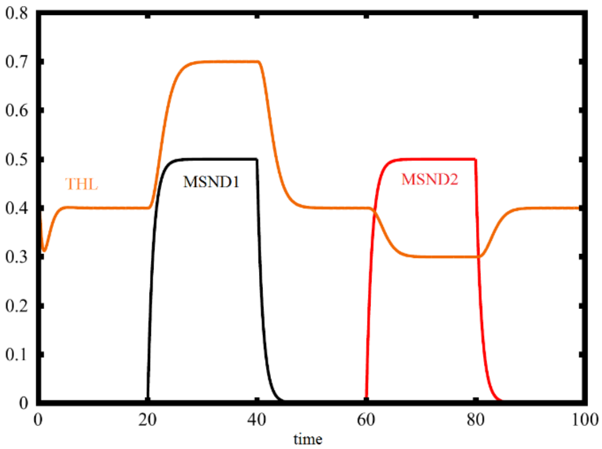


Figure 1. MSND1, MSND2 and THL activity. Mass model results show similar results spiking neural network. THL activity is effected by DA level similar to spiking neural network firing rate.

changing the parameter DA in **Equation 6**. The simulation interval is different as the it is scaled for the mass model, but the results shown in **Figure 4**, resembles the change in firing rates given in **Figure 3**.

4.3.Discussion of the results

As the simulation results given in Subsections 4.1 and 4.2 reveal, the overall activity of neuronal population can be followed from the mass model proposed as a simple linear system. Of course, with this coarse linear system approach, it is not possible to investigate the synaptic activity, or the activity in a neuronal population, but a general idea about the effect of dopamine can be followed easily. So building a simple model which consolidate the afferent and efferent connections between neuronal populations and the effect of neurotransmitter, would be informative to grasp the dynamics behind the neuronal activity. Furthermore, such a simple model can be versatile for the implementation of biologically inspired approaches and developing new learning rules as in [4,5].

5. References

[1] S. A. Prescott, S. Ratté, Y. De Koninck, and T. J. Sejnowski, "Pyramidal neurons switch from integrators in vitro to resonators under in vivo-like conditions," *Journal of neurophysiology*, vol. 100, no. 6, pp. 3030--3042, 2008.

[2] D. Terman, J. E. Rubin, A. C. Yew, and C. J. Wilson, "Activity patterns in a model for the subthalamopallidal network of the basal ganglia," *The Journal of Neuroscience*, vol. 22, no. 7, pp. 2963--2976, 2002.

[3] G. Deco, V. K. Jirsa, P. A. Robinson, M. Breakspear, and K. Friston, "The dynamic brain: From spiking neurons to neural masses and cortical fields," *PLoS Comput Biol*, vol. 4, no. 8, p. e1000092, 08 2008.

[4] E. Erçelik, R. Elibol, and N. S. Şengör, "A model on building and modifying the stimulus action association in the brain," in *2015 23nd Signal Processing and Communications Applications Conference (SIU)*, pp. 2533—2536, May 2015.

[5] E. Erçelik and N. S. Şengör, "A neurocomputational model implemented on humanoid robot for learning action selection," in *2015 International Joint Conference on Neural Networks (IJCNN)*, pp.1—6, July 2015.

[6] S. B. Furber, F. Galluppi, S. Temple, and L. A. Plana, "The spinnaker project," *Proceedings of the IEEE*, vol. 102, no. 5, pp. 652--665, May 2014.

[7] B. V. Benjamin, P. Gao, E. McQuinn, S. Choudhary, A. R. Chandrasekaran, J. M. Bussat, R. Alvarez-Icaza, J. V. Arthur, P. A. Merolla, and K. Boahen, "Neurogrid: A mixed-analog-digital multichip system for large-scale neural simulations," *Proceedings of the IEEE*, vol. 102, no. 5, pp. 699--716, May 2014.

[8] M. Dipoppa and B. S. Gutkin, "Flexible frequency control of cortical oscillations enables computations required for working memory," *Proceedings of the National Academy of Sciences*, vol. 110, no. 31, pp. 12\,828--12\,833, 2013.

[9] K. Gurney, T. J. Prescott, and P. Redgrave, "A computational model of action selection in the basal ganglia. i. a new functional anatomy," *Biological Cybernetics*, vol. 84, no. 6, pp. 401--410, 2001.

[10] E. M. Izhikevich and G. M. Edelman, "Large-scale model of mammalian thalamocortical systems," *Proceedings of the national academy of sciences*, vol. 105, no. 9, pp. 3593--3598, 2008.

[11] C. Yucelgen, B. Denizdurduran, S. Metin, R. Elibol, and N. S. Sengor, "A Biophysical Network Model Displaying the Role of Basal Ganglia Pathways in Action Selection", in *Artificial Neural Networks and Machine Learning -- ICANN 2012: 22nd International Conference on Artificial Neural Networks, Lausanne, Switzerland, September 11-14, 2012, Proceedings, Part I*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 177—184, 2012. doi="10.1007/978-3-642-33269-2_23"

[12] K. Doya, "What are the computations of the cerebellum, the basal ganglia and the cerebral cortex?" *Neural Networks*, vol. 12, no. 7–8, pp. 961 --974, 1999.

[13] J. G. Taylor and N. R. Taylor, "Analysis of recurrent cortico-basal ganglia-thalamic loops for working memory," *Biological Cybernetics*, vol. 82, no. 5, pp. 415--432, 2000.

[14] K. Gurney, T. J. Prescott, and P. Redgrave, "A computational model of action selection in the basal ganglia. ii. analysis and simulation of behaviour," *Biological Cybernetics*, vol. 84, no. 6, pp. 411--423, 2001.

[15] M. D. Humphries, R. D. Stewart, and K. N. Gurney, "A physiologically plausible model of action selection and oscillatory activity in the basal ganglia," *The Journal of Neuroscience*, vol. 26, no. 50, pp. 12\,921--12\,942, 2006.

[16] M. Guthrie, C. Myers, and M. Gluck, "A neurocomputational model of tonic and phasic dopamine in action selection: A comparison with cognitive deficits in parkinson's disease," *Behavioural Brain Research*, vol. 200, no. 1, pp. 48 -- 59, 2009.

[17] B. W. Balleine, M. R. Delgado, and O. Hikosaka, "The role of the dorsal striatum in reward and decision-making," *Journal of Neuroscience* vol. 27, no. 31, pp. 8161--8165, 2007.

[18] Q. J. Huys, A. Beck, P. Dayan, and A. Heinz, "Neurobiological understanding and computational models of addictive behavior," *Phenomenological Neuropsychiatry: Bridging the Clinic and Clinical Neuroscience: Springer: Science*}, 2011.

[19] D. Belin, S. Jonkman, A. Dickinson, T. W. Robbins, and B. J. Everitt, "Parallel and interactive learning processes within the basal ganglia: Relevance for the understanding of addiction," *Behavioural Brain Research*, vol. 199, no. 1, pp. 89 -- 102, 2009.
[20] G. Alexander, M. Crutcher, and M. DeLong, "Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions," *Progress in brain research*}, vol. 85, pp. 119--146, 1990.
[21] S. N. Haber, J. L. Fudge, and N. R. McFarland, "Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum," *The Journal of Neuroscience*, vol. 20, no. 6, pp. 2369--2382, 2000.
[22] S. N. Haber and B. Knutson, "The reward circuit: Linking primate anatomy and human imaging," *Neuropsychopharmacology*, vol. 35, no. 1, pp. 4 -- 26, 2010.
[23] G. S. Berns and T. J. Sejnowski, "How the Basal Ganglia Make Decisions".in *Neurobiology of Decision-Making* Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 101—113, 1996.
[24] M. R. DeLong and T. Wichmann, "Circuits and circuit disorders of the basal ganglia," *Archives of Neurology*, vol. 64, no. 1, pp. 20--24, 2007.
[25] M. H. Teicher, C. M. Anderson, A. Polcari, C. A. Glod, L. C. Maas, and P. F. Renshaw, "Functional deficits in basal ganglia of children with attention-deficit/hyperactivity disorder shown with functional magnetic resonance imaging relaxometry," *Nature Medicine*, vol. 6, no. 4, pp. 470 -- 473, 2000.
[26] R. C. Chacko, M. A. Corbin, and R. G. Harper, ``Acquired obsessive-compulsive disorder associated with basal ganglia lesions," *The Journal of Neuropsychiatry and Clinical Neurosciences*, vol. 12, no. 2, pp. 269--272, 2000.
[27] M. G. Murer, K. Y. Tseng, F. Kasanetz, M. Belluscio, and L. A. Riquelme, "Brain oscillations, medium spiny neurons, and dopamine," *Cellular and Molecular Neurobiology*, vol. 22, no. 5, pp. 611--632, 2002.
[28] A. L. Benabid, "Deep brain stimulation for parkinson's disease," *Current Opinion in Neurobiology*, vol. 13, no. 6, pp. 696 -- 706, 2003.
[29] C. D. Cunha, S. L. Boschen, A. Gómez-A, E. K. Ross, W. S. Gibson, H.-K. Min, K. H. Lee, and C. D. Blaha, "Toward sophisticated basal ganglia neuromodulation: Review on basal ganglia deep brain stimulation," *Neuroscience \& Biobehavioral Reviews*, vol. 58, pp. 186 -- 210, 2015.
[30] J. Sui, G. D. Pearlson, Y. Du, Q. Yu, T. R. Jones, J. Chen, T. Jiang, J. Bustillo, and V. D. Calhoun, "In search of multimodal neuroimaging biomarkers of cognitive deficits in schizophrenia," *Biological Psychiatry*, vol. 78, no. 11, pp. 794 -- 804, 2015.
[31] L. A. Gunaydin and A. C. Kreitzer, "Cortico-Basal Ganglia Circuit Function in Psychiatric Disease," in *Annual Review of Physiology,* vol. 78, pp. 327--350 2016.
[32] H. S. Mayberg, A. M. Lozano, V. Voon, H. E. McNeely, D. Seminowicz, C. Hamani, J. M. Schwalb, and S. H. Kennedy, "Deep brain stimulation for treatment-resistant depression," *Neuron,* vol. 45, no. 5, pp. 651 --660, 2005.
[33] B. H. Kopell and B. D. Greenberg, "Anatomy and physiology of the basal ganglia: Implications for DBS in psychiatry," *Neuroscience & Biobehavioral Reviews*, vol. 32, no. 3, pp. 408 -- 422, 2008.
[34] W. Schultz, P. Dayan, and P. R. Montague, "A neural substrate of prediction and reward," *Science*, vol. 275, no. 5306, pp. 1593--1599, 1997.
[35] R. E. Suri and W. Schultz, "Learning of sequential movements by neural network model with dopamine-like reinforcement signal, "*Experimental Brain Research,* vol. 121, no. 3, pp. 350--354, 1998.
[36] R. Suri, J. Bargas, and M. Arbib, "Modeling functions of striatal dopamine modulation in learning and planning," *Neuroscience*, vol. 103, no. 1, pp. 65 -- 85, 2001.
[37] P. Dayan and B. W. Balleine, "Reward, motivation, and reinforcement learning," *Neuron*, vol. 36, no. 2, pp. 285 -- 298, 2002.
[38] D. Joel, Y. Niv, and E. Ruppin, "Actor–critic models of the basal ganglia: new anatomical and computational perspectives," *Neural Networks*, vol. 15, no. 4–6, pp. 535 -- 547, 2002.
[39] P. R. Montague, S. E. Hyman, and J. D. Cohen, "Computational roles for dopamine in behavioural control," *Nature*, vol. 431, pp. 760 -- 767, 2004.
[40] M. Haruno and M. Kawato, "Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops: fMRI examination in stimulus-action-reward association learning," *Neural Networks*}, vol. 19, no. 8, pp. 1242 -- 1254, 2006.
[41] G. S. Berns and T. J. Sejnowski, "A model of basal ganglia function unifying reinforcement learning and action selection," in *Joint Symposium on Neural Computation*}, 1994.
[42] N. Şengör, O. Karabacak, and U. Steinmetz, "A computational model of cortico-striato-thalamic circuits in goal-directed behaviour," in *Artificial Neural Networks - ICANN 2008*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, vol. 5164, pp. 328—337, 2008.
[43] B. S. Gutkin, S. Dehaene, and J.-P. Changeux, "A neurocomputational hypothesis for nicotine addiction," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 4, pp. 1106--1111, 2006.
[44] S. H. Ahmed, M. Graupner, and B. Gutkin, "Computational approaches to the neurobiology of drug addiction," *Pharmacopsychiatry*, vol. 42, no. S144, pp. 144--152, 2009.
[45] S. Metin and N. S. Sengor, "From occasional choices to inevitable musts: A computational model of nicotine addiction," *Intell. Neuroscience*, vol. 2012, pp. 18:18--18:18, Jan. 2012.
[46] M. M. McCarthy, C. Moore-Kochlacs, X. Gu, E. S. Boyden, X. Han, and N. Kopell, "Striatal origin of the pathologic beta oscillations in Parkinson's disease," *Proceedings of the National Academy of Sciences*, vol. 108, no. 28, pp. 11\,620--11\,625, 2011.
[47] L. Squire, D. Berg, F. Bloom, S. du Lac, A. Ghosh, and N. Spitzer, *Fundamental Neuroscience*, 3rd ed., Elsevier, 2008.
[48] E. R. Kandel, J. H. Schwartz, T. M. Jessell, and S. Mack, *Principles of neural science*, New York, Chicago, San Francisco: McGraw-Hill Medical, 2013.
[49] M. J. Frank, "Hold your horses: A dynamic computational role for the subthalamic nucleus in decision making," in *Neural Networks*, vol. 19, no. 8, pp. 1120 -- 1136, 2006.
[50] A. Rosell and J. M. Giménez-Amaya, "Anatomical re-evaluation of the corticostriatal projections to the caudate nucleus: a retrograde labeling study in the cat," in *Neuroscience Research*, vol. 34, no. 4, pp. 257 --269, 1999.
[51] E. Y. Choi, B. T. T. Yeo, and R. L. Buckner, "The organization of the human striatum estimated by intrinsic functional connectivity," *Journal of Neurophysiology*, vol. 108, no. 8, pp. 2242--2263, 2012.

[52] R. N. Cardinal, C. A. Winstanley, T. W. Robbins, and B. J. Everitt, "Limbic corticostriatal systems and delayed reinforcement," *Annals of the New York Academy of Sciences*}, vol. 1021, no. 1, pp. 33--50, 2004.
[53] G. Alexander and M. Crutcher, "Functional architecture of basal ganglia circuits: neural substrates of parallel processing," *Trends Neurosci*., vol. 13, no. 7, pp. 266--271, 1990.
[54] S. Sakai, "Cerebellar thalamic and thalamocortical projections," in *Handbook of the Cerebellum and Cerebellar Disorders*, M. Manto, J. Schmahmann, F. Rossi, D. Gruol, and N. Koibuchi, Eds. Springer Netherlands, pp.529—547, 2013.
[55] P. van Donkelaar, J. F. Stein, R. E. Passingham, and R. C. Miall, "Temporary inactivation in the primate motor thalamus during visually triggered and internally generated limb movements," *Journal of Neurophysiology*, vol. 83, no. 5, pp. 2780--2790, 2000.
[56] B. Denizdurduran and N. S. Şengör, "A realization of goal-directed behavior-implementing a robot model based on cortico-striato-thalamic circuits." in *ICAART*, pp. 289—294,2012.
[57] O. Karabacak and N. S. Sengor, "A Computational Model for the Effect of Dopamine on Action Selection During Stroop Test", İn: *Kollias S.D., Stafylopatis A., Duch W., Oja E. (eds) Artificial Neural Networks – ICANN 2006. ICANN 2006. Lecture Notes in Computer Science*, vol 4131, pp. 485--494 Springer, Berlin, Heidelberg, 2006.
[58] E. Izhikevich, "Simple model of spiking neurons," *Neural Networks, IEEE Transactions on*}, vol. 14, no. 6, pp. 1569--1572, Nov 2003.
[59] D. Goodman and R. Brette, "Brian: a simulator for spiking neural networks in python," *Frontiers in Neuroinformatics*, vol. 2, p. 5, 2008.
[60] R. Elibol and N. S. Şengör, "A computational model to investigate the effect of dopamine on neural synchronization in striatum,"*2015 International Joint Conference on Neural Networks (IJCNN)*, Killarney, pp. 1-5, 2015. doi: 10.1109/IJCNN.2015.7280762
[61] R. Elibol and N. S. Şengör, "A Computational Model of Neural Synchronization in Striatum.", in *Computational Neurology and Psychiatry* Volume 6 of the series Springer Series in Bio-/Neuroinformatics pp 373-403, 2017. doi: 10.1007/978-3-319-49959-8_13.
[62] R. Elibol and N. S. Şengör, "Looking at the role of direct and indirect pathways in basal ganglia networks at different levels," *BMC Neuroscience*, vol. 16, no. 1, p. P225, 2015. doi: 10.1186/1471-2202-16-S1-P225.
[63] R. Elibol and N. S. Şengör, "Modeling basal ganglia circuits with mass model equations,"*2016 Medical Technologies National Congress (TIPTEKNO)*, Antalya, Turkey, 2016. doi: 10.1109/TIPTEKNO.2016.7863131.

Rahmi Elibol received B. Sc. degrees from Karadeniz Technical University, Turkey, in 2004, and M. Sc. degrees from Istanbul Technical University, Turkey, in 2013. He joined Istanbul Technical University as research assistant in 2011, where he is a PhD student now. His research interests include nonlinear dynamical systems, mathematical modeling and computational neuroscience

Neslihan Serap Şengör received B. Sc., M. Sc. And Ph.D. degrees from Istanbul Technical University, Turkey, in 1985, 1988, 1995 respectively. She joined Istanbul Technical University as research assistant in 1986, where she is Professor now. She worked as visiting scientist in Helsinki University of Technology 2000-2001 and summer of 2006. She also visited Lincoln University, UK as research associate in the summer of 2015. Her research interests include nonlinear circuits and systems, neural networks.

EMOTION RECOGNITION VIA GALVANIC SKIN RESPONSE: COMPARISON OF MACHINE LEARNING ALGORITHMS AND FEATURE EXTRACTION METHODS

Deger AYATA[1], Yusuf YASLAN[1], Mustafa KAMASAK[1]

[1] Faculty of Computer and Informatics Engineering, Istanbul Technical University, Istanbul, Turkey
{ayatadeger, yyaslan, kamasak}@itu.edu.tr

Abstract: *Emotions play a significant and powerful role in everyday life of human beings. Developing algorithms for computers to recognize an emotional expression is widely studied area. In this study, emotion recognition from Galvanic Skin Response signals was performed using time domain, wavelet and Empirical Mode Decomposition based features. Valence and arousal have been categorized and relationship between physiological signals and arousal and valence has been studied using k-Nearest Neighbors, Decision Tree, Random Forest and Support Vector Machine algorithms. We have achieved 81.81% and 89.29% accuracy rate for arousal and valence respectively.*
Keywords: *Biomedical Signal Processing, Emotion Recognition, Pattern Recognition, Machine Learning, Physiological Signal, Galvanic Skin Response, Decision Tree, Random Forest, k-Nearest Neighbors, Support Vector Machine.*

1. Introduction

Emotions play a significant and powerful role in everyday life of human beings. The importance of emotions motivated the researchers in the biomedical engineering, computer and electronics engineering disciplines to develop automatic methods for computers to recognize emotional expressions [1]. For a rich set of applications including human-robot interaction, computer aided tutoring, emotion aware interactive games, neuro marketing, socially intelligent software apps, computers should consider the emotions of their human conversation partners. Speech analytics and facial expressions have been used for emotion detection. Ekman et al. stated that six different facial expressions (fearful, angry, sad, disgust, happy, and surprise) were categorically recognized by humans from distinct cultures using a standardized stimulus set [2]. However, using only speech signals or facial expression signals have disadvantages: using only them is not reliable to detect emotion, especially when people want to conceal their feelings. Compared with facial expression, using physiological signals is a reliable approach to probe the internal cognitive and emotional changes of users.

In this study, emotion recognition from Galvanic Skin Response (GSR) was performed using time domain based features, wavelet approaches and Empirical Mode Decomposition (EMD) approaches. The study compares machine learning algorithms and feature extraction methods for GSR based emotion recognition. Valence and arousal have been categorized and relationship between physiological signals and arousal and valence has been studied using Decision Tree (DT), Random Forest (RF), k-Nearest Neighbors (kNN) and Support Vector Machine (SVM) learning algorithms.

We have achieved 81.81% and 89.29% accuracy rate for arousal and valence respectively by using only Galvanic Skin Response signal. We have also showed that using convolution has positive effect on accuracy rate compared to non-overlapping window based feature extraction.

The outline of the paper is as follows. Section 2 summarizes related work about emotion recognition with GSR. Section 3 describes methods in detail, including emotion representation, data collection, data preprocessing, feature extraction and classification. Results are presented and discussed in Section 4. The paper ends with a conclusion in Section 5.

2. Related Work

Emotions regulate the autonomic nervous system, which, in turn, causes variations in the secretion of sweat on the skin's surface, as well as changes in the heart rate and respiration rate [3].

GSR, which is known also as Electro Dermal Activity (EDA) is a low cost, easily captured physiological signal. GSR is a reflection of physiological reactions that generate excitement. Emotional arousal induces a sweat reaction, which is particularly prevalent at the surface of the hands and fingers and the soles of the feet. When people get excited, body sweats, the amount of salt in the skin increases and the skin's electrical resistance also increases.

Received on: 30.01.2017
Accepted on: 15.03.2017

GSR appears sensitive only to the arousal dimension not direction or valence of the emotion involved. Skin conductivity varies with changes in skin moisture level(sweating) and can reveal changes in sympathetic nervous system. Nakasone et al. have used skin conductance and muscle activity for emotion recognition [4]. Nourbakhsh et al. investigated different time and frequency domain features of GSR in multiple difficulty levels of arithmetic and reading experiments [5]. Channel et al. has conducted a research on emotion assessment related to arousal evaluation using EEG's and peripheral physiological signals. They have used Galvanic Skin Resistance (GSR), blood pressure, temperature as well as EEG data. They have used Naïve Bayes and Fisher Discriminant Analysis (FDA) classifiers [6].

In this study we have used DT, RF, k-NN and SVM learning algorithms using time domain based features, wavelet and EMD approaches.

3. Materials and Methods

Biosensors can monitor physiological attributes of the human body that are controlled directly by autonomic nervous system. These sensors can collect signals including skin conductance, blood volume, temperature, heart rate. Physiological data is challenging to represent and process due to its noise, volume and multimodality. Moreover a persons' emotional response may be different from another.

3.1. Emotion Recognition Using GSR

In this study, we have used Galvanic Skin Response Signals. GSR, which is known also as EDA is a low cost, easily captured physiological signal [4,5,6]. GSR is a reflection of physiological reactions that generate excitement. Skin reacts when it is exposed to emotionally loaded images, videos, events, or other kinds of stimuli, no matter if it is positive or negative. Emotional changes induces a sweat reaction, which is particularly prevalent at the surface of the hands and fingers and the soles of the feet. When people get excited, body sweats, the amount of salt in the skin increases and the skin's electrical resistance also increases. Change in emotions trigger the sweat glands in our body, and make them more active. Whenever sweat glands become more active, they secrete moisture towards the skin surface. That changes the balance of positive and negative ions and affects the electrical currents' flow property on skin and it is most observable on hands and feet. This resistance decreases due to an increase of perspiration, which usually occurs when one is experiencing emotions such as stress or surprise. Resulting changes in skin conductance are measurable and generally termed as Galvanic Skin Response. In GSR method, the electrical conductance of the skin is measured through one or two sensor(s) usually attached to hand or foot.

If the subject's hands are static, like when passively watching a video, then the recommended recording locations are index and middle fingers. In case of the subjects use their both hands, like when using a keyboard and a mouse, then the recommended recording locations are hand palms. However, if the subjects use their both hands, but quite extensively, like when manipulating and interacting with real-life environments, then the recommended recording locations are foot soles. Sensors should be used in inner sides so as not to be affected by the pressure while standing or walking.

In our study, GSR signals have been captured from left hand fingers.

3.2. Emotion Representation

Psychologists proposed and identified different models for representing emotions. There are two significantly different models for representing emotions: the categorical model and the dimensional model. The categorical model and dimensional models have two different methods for estimating the actual emotional states of a person. In the categorical model emotions are labelled. The person is "happy" or "sad" and people get a sense of what is meant. In the dimensional model the representation is based on a set of quantitative measures using multidimensional scaling (e.g. "pleasant-unpleasant") [2,7,8].

The emotion valence-arousal dimensional model, represented in Figure 1, is widely used in many research studies. The Pleasure - Displeasure Scale measures how pleasant an emotion may be. Pleasure(Valence) ranges from unpleasant to pleasant and it is the degree of attraction of a person toward a specific object or event. It ranges from negative to positive. The Arousal-Non Arousal Scale measures the intensity of the emotion. The arousal is a physiological and psychological state of being awake or reactive to stimuli, ranging from passive to active [7,8].

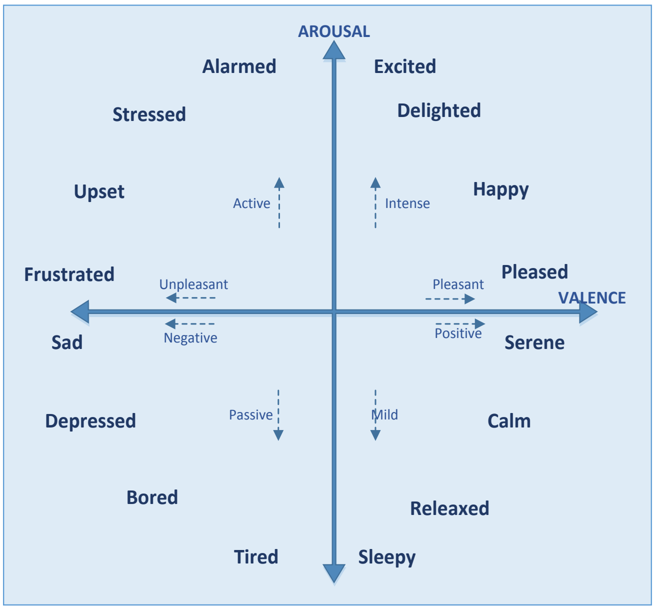


Figure 1. Valence – Arousal Model

Valence-arousal model chart is a model for emotions to be mapped out by range of arousal and valence that is experienced during a particular emotion. The Valence-axis and Arousal-axis separate the coordinate plane into four regions. Let α be the emotional state observed, in valance-arousal plane, a subject can be in one of emotion sets that can be described as follows:

$$\alpha \begin{cases} if\ valence > 0 \wedge arousal > 0,\ \alpha \in \{excited, delighted, happy, pleased, interested, convinced\} \\ if\ valence < 0\ \wedge arousal > 0,\ \alpha \in \{alarmed, stressed, upset, frustrated, insulted, hostile\} \\ if\ valence < 0\ \wedge arousal < 0,\ \alpha \in \{sad, depressed, bored, tired, worried, hesitant, doubtful\} \\ if\ valence > 0 \wedge arousal < 0,\ \alpha \in \{sleepy, relaxed, calm, serene, impressed, peaceful, confident\} \end{cases}$$

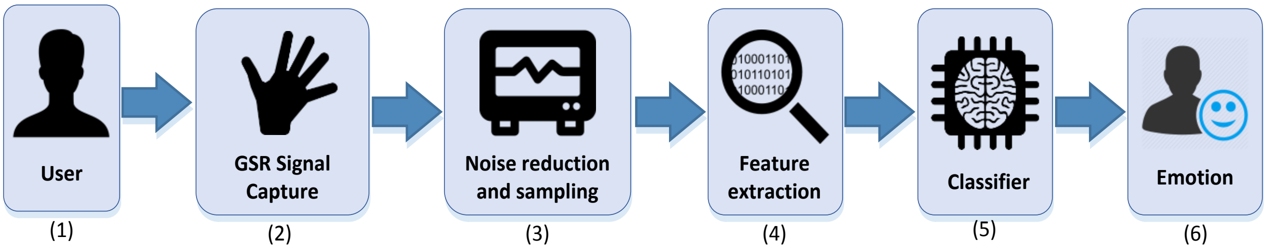


Figure 2. GSR Based Emotion Recognition Pipeline

Table 1. Basic Features and Formulas Used

Attribute	Formula	Attribute	Formula
Minimum	$\min[X_n]$	*Skewness*	$\sum_{n=1}^{N}(X_n - AM)\frac{3}{(N-1)SD^3}$
Maximum	***max[X$_n$]***	*Kurthosis*	$\sum_{n=1}^{N}(X_n - AM)\frac{4}{(N-1)SD^4}$
Arithmetic Mean (AM)	$\frac{1}{N}\sum_{n=1}^{N}X_n$	*Median*	$\frac{\left(\frac{N}{2}\right)^{th} \text{value} + \left(\frac{N}{2}+1\right)^{th} \text{value}}{2}$ or $\left(\frac{N+1}{2}\right)^{th}$
Mean Absolute	$\frac{1}{N}\sum_{n=1}^{N}\lvert X_n \rvert$	*Moment (k^{th} order)*	$\frac{1}{N}\sum_{n=1}^{N}{X_n}^k$
Root Mean Square	$\sqrt{\frac{1}{N}\sum_{n=1}^{N}(X_n)^2}$	*First Degree Difference*	$\frac{1}{N-1}\sum_{n=1}^{N}\lvert X_{n+1} - X_n \rvert$
Standard Deviation (SD)	$\sqrt{\frac{1}{N}\sum_{n=1}^{N}(X_n - AM)^2}$	*Second Degree Difference*	$\frac{1}{N-2}\sum_{n=1}^{N}\lvert X_{n+2} - X_n \rvert$

3.3. Emotion Recognition Pipeline

The pipeline we have used in this study is depicted in Figure 2. Galvanic Skin Response signal is captured from subjects through GSR biosensor (1,2). Noise reduction and sampling process is done (3). Feature extraction methods are applied to GSR signal(4), and results represented as feature vectors. Then, feature vectors are fed to classifier. Classifier takes this feature vector as input (5) and makes a prediction about the emotional state of the user (6) by estimating arousal and valence values.

3.4. Dataset

Deap is a multimodal dataset for the analysis of human affective states. In the dataset EEG and peripheral physiological signals of 32 participants were recorded as each watched 40 videos, each video is one-minute long excerpts of music videos. Music video clips are used as the visual stimuli to elicit different emotions.

Participants rated each video in terms of the levels of arousal, valence, like/dislike, dominance and familiarity. For 22 of the 32 participants, frontal face video was also recorded. The dataset was first presented by Kolestra et al. [9]. The data was downsampled to 128Hz, EOG artefacts were removed, a bandpass frequency filter from 4.0 - 45.0Hz was applied and, the data was segmented into 60 second trials and a 3 second pre-trial.

The total signal record time for each video is 63 second and sampling frequency is 128 Hz which means for each channel 8064 sample data points have been collected. The dataset contains both EEG and peripheral physiological signals. In this paper, among recorded signals Galvanic Skin Response signals have been considered. Galvanic skin response signals have been recorded from left hand middle and ring fingers.

3.5. Feature Extraction

Features from signals have been extracted in the time domain and based on statistics. Wavelet and Empirical Mode Decomposition approaches are also used during feature extraction process.

3.5.1. Time Domain Features

GSR signal has been subjected to various length moving windows for feature extraction. In each trial, we have obtained signals and divide each channel signal into segments (e.g. 20 segments with 3s length per segment). Features have been first extracted from each window, and their values across the consecutive windows have been concatenated for each subject and for each video.

In the time domain, arithmetic mean value, maximum value, minimum value, standard deviation, variance, skewness coefficient, kurtosis coefficient, median, number of zero crossings, entropy, mean energy, moments, change in signal values have been considered as features. Table 1 depicts feature list and formula of pertaining features.

In order to capture right attributes for emotion classification, various attributes have been selected as feature set and relationship between arousal and valence has been studied. Table 2 shows studied feature sets and their attributes. FS-10 which includes 10 basic attributes has been used as base set. By enriching FS-10 with different order of moments FS-14 has been obtained. FS-18 contains both FS 14 and for additional attributes. FS-22 is the largest attribute set with 22 features.

Table 2. Feature Sets and Attributes

Feature Set (FS)	Attributes
FS-10	Minimum, Maximum, Arithmetic Mean (AM), Standard Deviation, Variance, Skewness, Kurtosis, Median, Zero Crossings, Mean Energy
FS- 14	Feature 10 Set, 3^{rd}, 4^{th}, 5^{th} , 6^{th} Moment
FS- 18	Feature 14 Set, Mean Absolute Value, Max Scatter Difference, Root Mean Square, Mean Absolute Deviation
FS - 22	Feature 18 Set, 1^{st} Degree Difference, 2^{nd} Degree Difference, 1^{st} Degree Diff Divided with Std Deviation, 2^{nd} Degree Diff Divided with Std Deviation

3.5.2. Discrete Wavelet Transformation

Since biological signals are non-stationary and changes over time in nature, Fourier transformation is inconvenient to analyze GSR signals. GSR signals are not periodic and their amplitude, phase and frequencies change. Wavelet transformation is generally can deal with non-stationary signals. Discrete Wavelet Transformation is a method developed to overcome the deficiencies of the Fourier transformation over non-stationary signals and this method is less sensitive towards noise and can be easily applied to non-stationary signals [10]. Features have been extracted using Discrete Wavelet Transform. For the DWT, it is important to identify appropriate wavelet type and determining the level of decomposition. Daubechies db2 has been selected as wavelet.

The features are the sum of absolute amplitudes, min, max, mean energy, sum of squares, kurtosis, skewness, and standard deviation.

3.5.3. Empirical Mode Decomposition

In this study, we proposed and evaluated the use of Empirical Mode Decomposition (EMD) technique. The

GSR signal data were separated into intrinsic mode functions (IMFs) using the EMD method. EMD is the fundamental part of the Hilbert-Huang transform (HHT) which is a way to decompose a signal into so-called intrinsic mode functions (IMF) along with a trend, and obtain instantaneous frequency data [11, 12]. Empirical mode decomposition (EMD) and the Hilbert spectral analysis (HSA) are used together in HHT and act as a signal transform method. But EMD can be used separately as a signal feature extraction method, too. It is used in a variety of studies such as decomposition of speech signal [13], epileptic seizure detection in EEG signals [14], extraction of significant features [15] etc.

The algorithm itself depends on enveloping the signal functions maxima and minima, finding the mean, extracting an IMF and iterate this steps until the peak frequency becomes smaller than the defect one. In our study we have used EMD for feature extraction.
After applying EMD, we have extracted Minimum, Maximum, Average, Standard Deviation, Variance, Skewness, Kurtosis, Median, Zero Crossings, Mean Energy, 3rd Moment, 4th Moment, 5th Moment and 6th Moment as features.

3.6. Classification

Labeling the samples is critical for Machine Learning. Arousal and Valence values have been categorized to two (Low, High) classes. We divide the trials into classes according to each trial's rating value (high: ≥ 4.5, low: < 4.5). GSR signals taken from 32 subjects all have been used for training and test steps. After feature extraction the signals are classified into classes using classifiers Decision Tree(J48), Random Forest, Support Vector Machine (SVM) and k-Nearest Neighbors(kNN).

A decision tree is a non-parametric supervised learning method that predicts the value of a target variable by learning decision rules from the data and used for classification and regression. Decision tree partitions dataset into groups as homogeneous as possible in terms of the variable to be predicted. Attribute selection is the fundamental step to construct a decision tree. Entropy and Information Gain is used to process attribute selection. ID3 and C4.5(aka J48) algorithms have been introduced by J.R Quinlan which produce reasonable decision trees. C4.5 is an extension of ID3 algorithm [16].

Random Forests are an ensemble method with which classification and regression are performed using a forest of decision trees, each constructed using a random subset of the features. Random forests achieve high accuracy in a variety of problems, making them versatile choice for many applications. Since only a subset of the features used, random forests capable of handling high dimensional data. Also, a trained model can be used to determine the pairwise proximity between samples. These features make random forests a popular technique in bioinformatics and specialized random forests for these purposes are an active area of research [17].

The support vector machine (SVM) is a supervised method that constructs a hyperplane separating groups based on a set of given training data in a multidimensional space. Objective of the SVM is to find the optimal separating hyperplane which maximizes the margin of the training data. SVM supports both regression and classification tasks. SVMs can perform linear classification tasks. SVMs can also perform a non-linear classification using what is called the kernel trick, by mapping their inputs implicitly into high-dimensional feature spaces. SVMs produce robust, accurate predictions, and are least affected by noisy data, and are less prone to overfitting [18].

k-Nearest Neighbors algorithm (kNN) is a nonparametric method used for classification and regression [19]. C4.5 builds a decision tree classification model during training. SVM builds a hyperplane classification model during training. kNN does not build such classification model, it just stores the labeled training data. For a new unlabeled instance, it looks at the k - closest labeled training data points and then using the neighbors' classes and determines class.

4. Experimental Results and Discussion

We have conducted tests with time domain only features, wavelet and EMD approaches. During test process, we have tested all feature sets and compared the results in time and frequency domains. We have used feature sets as part of feature vectors to train and test classifiers. In order to compare classifier performances we have also conducted test cases with Decision Tree(J48), Random Forest, Nearest Neighbors (kNN) classifiers separately.

4.1. Time based Statistical Features Tests

Tests have been conducted with 10-fold cross validation by using Random Forest machine learning algorithm. Window Duration W ∈ {1, 3, 5, 8, 10,12,15, 30, 60} s, Feature Set Size FS ∈ {10, 14, 18, 22}, and Convolution C ∈ {Convolution, Non-Convolution} setups have been tested with various combinations.

Window Duration Size Tests

Window duration has effect on accuracy rate. Various window size duration between 1 seconds and 60 seconds have been selected. Tests with 3 seconds window duration performed better than other window duration size. Results are depicted in Table 3 and Figure 3.

Feature Set Tests

Feature extraction has effect on accuracy rate. Various feature sets(FS) have been selected. Tests with FS 10, FS 14, FS 18 and FS 22 has been conducted. FS 14 performed better than other feature sets, corresponding results are depicted in Figure 4.

Convolution vs. Non-Convolution Tests

Windows have been slided by collapse or not collapse manner. Overlapped and one second slide duration has performed better compared to non-overlapping window sliding. Figure 3 and Figure 4 confirms that convolution is generally a better approach to increase accuracy rate.

Table 3. GSR – Time Domain Statistics Experiments

Feature Size	Record Size	Class Size	AROUSAL Accuracy No -Conv	AROUSAL Accuracy Convolution	VALENCE Accuracy No-Conv	VALENCE Accuracy Convolution	Window Duration (sn)
10x63	40x32	2	70.78%	70.78%	69.6%	69.6%	1
10x21	40x32	2	71.53%	71.46%	70.54%	71.04%	3
10x12	40x32	2	70.46%	70.76%	69.68%	70.39%	5
10x8	40x32	2	69.6%	70.17%	69.49%	70.23%	8
10x6	40x32	2	69.0%	70.15%	69.32%	69.76%	10
10x5	40x32	2	68.96%	69.45%	69.21%	69.56%	12
10x4	40x32	2	68.75%	68.9%	68.92%	69.07%	15
10x2	40x32	2	68.04%	68.44%	68.21%	68.37%	30
10x1	40x32	2	66.48%	66.48%	65.7%	65.7%	60

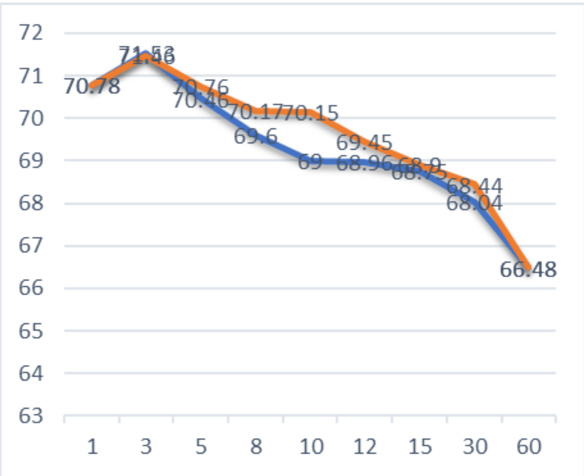


Figure 3. Windows Sizes vs.Accuracy Rate

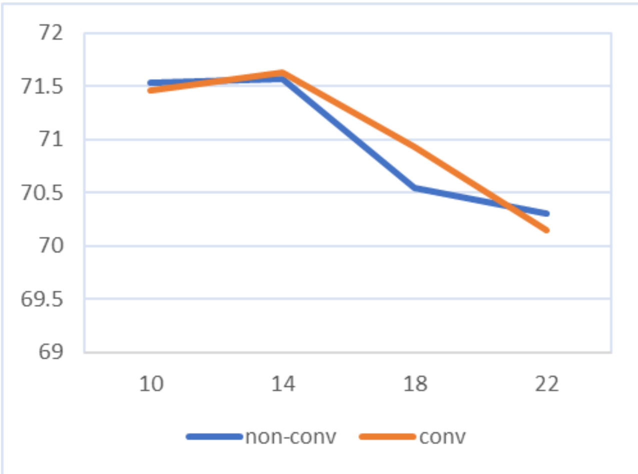


Figure 4. Feature Sets vs. Accuracy Rate

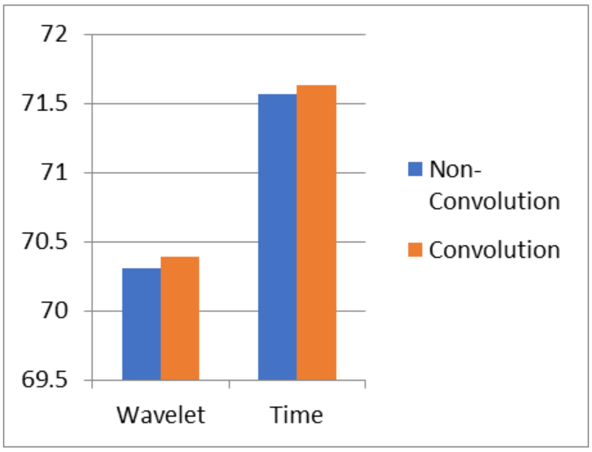


Figure 5. Wavelet and Time based features vs Accuracy

Table 4. Wavelet versusTime Domain Statistics Experiments

Class	**Wavelet**		**Time**	
Arousal	70.31%	70.39%	71.53%	71.46%
Valence	70.7%	70.31%	70.54%	71.04%
	Non-Conv	Conv	**Non-**Conv	Conv

4.2. Wavelet versus Time Based Features Tests

Time based features and wavelet approach have been compared by tests. Time based features performed better as shown in Table 4 and Figure 5.

4.3. Classifier Comparison Tests

Tests have been conducted with 10-fold cross validation by using various classifiers C ∈ {Decision Tree(J48), Random Forest, Nearest Neighbors (kNN)}and Feature Set Size F ∈ {10} and Window Duration W∈ {3} seconds configurations. Table 5 depicts accuracy rates for various Classifiers for arousal and valence respectively.

Table 5. Accuracy Rates for various classifiers

Dimension	kNN	DT	RF	SVM
Arousal	58.12	59.21	71.53	71.40
Valence	60.54	59.20	71.04	70.54

4.4. EMD Features Tests

GSR signals have been tested with Window Duration W ∈ {3} s, Feature Set Size ∈ {FS-14}, since these setups were best with statistical only feature extraction methods. After feature extraction step, all features have been used as input vector to Random Forest classifier. To verify the effectiveness of this method, 32 subjects were tested.

Applying EMD for feature extraction gave better results both for arousal and valence dimensions. EMD performed better compared to time-only statistical feature extraction. The accuracy rate increased from 71.93% to 85.07% for arousal and from 71.04% to 82.81% for valence as depicted in Table 6 respectively.

Table 6. EMD based Results

Dimension	Non – EMD Accuracy %	EMD Based Accuracy %
Arousal	71.53	81.81
Valence	71.04	89.29

5. Conclusion and Future Work

The methods of recognizing arousal and valence values directly from only GSR Signals is a challenge task. In this study, an emotion recognition system based on GSR is introduced by considering affective and physiological computing approaches. Emotion recognition from GSR signals was performed. In this work, Valence and arousal have been categorized and relationship between GSR signals, arousal and valence has been studied using Decision Tree, Random Forest, k-Nearest Neighbor and Support Vector Machine algorithms.

We have seen that there is a relationship between GSR signals, arousal and valence. In case of we categorize both arousal and valence into two classes; we have achieved 71.53% and 71.04% accuracy rate for arousal and valence respectively with time domain only features.

Applying EMD increased accuracy rate both for arousal and valence dimensions. EMD performed better and yielded a modest increase in the performance compared to time-only statistical feature extraction. Based on the results, the accuracy rate increased from 71.93% to 85.07% for arousal and from 71.04% to 82.81% for valence. The results suggest that the proposed EMD based approach is effective for GSR signals, and EMD based feature extraction is worth for the further application in the physiological signal analysis.

For future works, we are planning to apply data fusion techniques with other physiological signals and apply different machine learning algorithms to increase accuracy rate.

6. References

[1] N. Sebe, I.Cohen, and T. S. Huang, *"Multimodal Emotion Recognition"*, WSPC, June 18, 2004

[2] P. Ekman, R.W.Levenson, , W.V. Friesen, "*Autonomic nervous system activity distinguishing among emotions"*, Science 221, 1208– 1210., 1983

[3] Shimmer, "*Measuring Emotion: Reactions To Media*", Dublin, Ireland, 2015

[4] A. Nakasone, H.Predinger, and M.Ishizuka, *"Emotion Recognition from Electromyography and Skin Conductance"*, in proc. BSI 2005(2005), 219-222

[5] N. Nourbakhsh, Y. Wang, F. Chen, R.. Calvo, "*Using Galvanic Skin Response for Cognitive Load Measurement in Arithmetic and Reading Tasks"*, ACM 2012

[6] G. Channel, J. Kronegg, D. Grandjean, T. Pun, *"Emotion Assessment : Arousal Evaluation Using EEG's and Peripheral Physiological Signals"*, Technical Report, Universite de Geneve, 2005.

[7] A.Mehrabian, (1980). "*Basic dimensions for a general psychological theory"*, pp. 39–53. ISBN 0-89946-004-6.

[8] A. Mehrabian; J. A. Russell (1974). "*An approach to environmental psychology (1 ed.)"*, Cambridge, Mass.: MIT Press.

[9] S. Koelstra, C. Muehl, M. Soleymani, A. Yazdani, J.-S. Lee, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, *"DEAP: A Database for Emotion Analysis Using Physiological Signals"*, IEEE Trans. Affective Computing, vol. 3, no. 1, pp. , Jan- Mar. 2012.

[10] A.N. Akansu, W.A. Serdijn, and I.W. Selesnick, "*Wavelet Transforms in Signal Processing: A Review of Emerging Applications"*, Physical Communication, Elsevier, vol. 3, issue 1, pp. 1–18, March 2010.

[11] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. Yen, C. C. Tung, H. H. Liu, *"The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis"*, The Royal Society, 1996.

[12] X. Yao, X. Sun, Y. Yang, D. Wu, X. Liang, *"Features extraction and reconstruction of country risk based on EMD"*, Procedia Computer Science, Volume 31, 2014, Pages 265-272.

[13] N. V. Davis, *"Feature extraction using empirical mode decomposition of speech signal"*, International Journa of

Engineering Trends and Technology, Volume 3, Issue 2, 2012.
[14] F.K. Onay, C. Köse, *"Epileptic seizure detection in EEG signals using recurrence plot of intrinsic mode functions"*, Eleco 2014, Bursa, Turkey.
[15] S. Cho, Y. Seo, *"Extraction of significant features using empirical mode decomposition and its application,* WCECS 2013, San Francisco, USA.
[16] Quinlan, J. R. C4.5: *Programs for Machine Learning*. Morgan Kaufmann Publishers, 1993.
[17] Ho, Tin Kam. *Random Decision Forests* . Procedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–28
[18] V.N. Vapnik, *The Nature of Statistical Learning Theory*, New York, NY: Springer, 2000
[19] N.Bhatia, V.Ashev, *"Survey of Nearest Neighbor Techniques"*, International Journal of Computer Science and Information Security, Vol. 8, No. 2, 2010.

Değer Ayata is PhDc at Computer Engineering Department, Istanbul Technical University. He got his B.Sc. degree in Electronic & Telecommunication Engineering from Istanbul Technical University, Turkey, in 1996. He got his M.Sc. degree in Computer and Control Engineering from ITU in 2000 and MBA degree from Bogaziçi University, Istanbul, Turkey in 2003. His research interests are signal processing, computer vision, natural language processing, machine learning, security, data mining, big data and artificial intelligence theory and applications.

Yusuf Yaslan is an Assistant Professor at Computer Engineering Department, Istanbul Technical University. He got his B.Sc. degree in Computer Science Engineering from Istanbul University, Turkey, in 2001. In 2002, he joined the Multimedia Signal Processing and Pattern Recognition laboratory at Istanbul Technical University (ITU). He got his M.Sc. degree in Telecommunication Engineering and his Ph.D. in Computer Engineering from ITU, in 2004 and 2011 respectively. During 2001 and 2002, he was an intern at the FGAN-FOM Research Institute, in Germany. He was a visiting researcher at Statistical Machine Learning and Bioinformatics Group, Aalto University in Finland during 2007 and 2008. His research interests are watermarking, machine learning theory and applications, especially in semi supervised learning, social networks, music recognition/recommendation and data mining applications on IoT data and smart cities. Currently, he is involved as member in the European FP7 Smart City project VITAL.

Mustafa E. Kamaşak received the B.S.E.E. and M.S.E.E. degrees from Bogazici University, Istanbul, Turkey, in 1997 and 1999, respectively, and the Ph.D. degree from Purdue University, West Lafayette, IN, in 2005. He is currently an Associate Professor with the Department of Computer Engineering, Istanbul Technical University, Istanbul, Turkey. His research interests include medical informatics, medical imaging, and image processing.

ESTIMATION OF OXYGEN SATURATION WITH LASER OPTICAL IMAGING METHOD

Arman JALALI PAHNVAR, Anıl ISIKHAN, Ibrahim AKKAYA*, Yusuf EFTELI, Mehmet ENGIN, Erkan Zeki ENGIN

Ege University, Electrical Electronics Engineering, Izmir, Turkey
{armanjalali35, aisikhan200}@gmail.com, {ibrahim.akkaya, yusuf.efteli mehmet.engin, erkan.zeki.engin}@ege.edu.tr

Abstract: *The aim of this study is to determine the estimation of hemoglobin concentration and oxygen saturation of tissue by non-invasively functional laser imaging for early skin cancer diagnosis. The early diagnosis of melanoma is a key factor that remarkably reduces the mortality rate. Diffuse reflectance spectroscopy is a very useful device for diagnosis and treatment purposes under in-vivo conditions. At this point, the aforementioned device, which takes into account the scattering of tissue, is to determine the concentration of chromophores (or optical absorbers) due to attenuated light strikes to the superficial layer of tissue. Laser-type light based imaging techniques in medical diagnosis substantially produce good results. Consequently, the aim of this study is to estimate* HbO_2 *(%) and Hb (%) concentrations by use of a single wavelength (680 nm).*

Keywords: *Melanoma; oxygen saturation; laser; non-invasive imaging.*

1. Introduction

The hemoglobin molecule in the red blood cells (RBC) serves to transport oxygen from the lungs to the tissues and binds to four oxygen atoms to form the oxygenated hemoglobin (HBO2) molecule [1]. Oxygen saturation (SO2) is the statistical mean of oxygenated hemoglobin (HBO2), depending on the total number of hemoglobin can be bound with the oxygen. Due to the fact that optical methods are non-invasive and allow the oxygen saturation to be continuously measured, the absorption of light by the blood due to oxygen saturation is an intensively studied subject [1 – 5]. Studies are concentrated at wavelength range between 250 – 1000 nm. Optical methods for measuring oxygen saturation in tissue are widespread in clinical conditions. These measurements are resulting from the absorption and scattering based losses along the path traveled by the light. Generally, the scattering losses are not often considered in data analysis.

On the other hand, in order to monitor the SO_2 value of whole circulation system of human, pulse oximetry is a non-invasive and widely used method [6, 7]. It depends on the light absorption changes of the arterial blood, while the absorption of the skin, muscle, bones and venous blood remain unchanged. Usually, the spectrophotometric method which uses multi-wavelength light source is used to monitor the arterial blood oxygen saturation [8, 9]. However, the accuracy of this method would be affected by opaque skin, irregular blood flow, motion of body, especially during hypo-perfusion. Moreover, it does not have the ability to monitor local changes of oxygen. The advantages of the proposed single-wavelength method contains less power consumption (only one wavelength needed), real-time monitoring (there is no need to switch wavelength) and simplicity (simple operational steps) for on-site and portable blood oxygenation monitoring applications.

Diffuse reflectance spectroscopy is a very useful device for diagnosis and treatment purposes under in-vivo conditions [10]. At this point, the aforementioned device, which takes into account the scattering of tissue, is to determine the concentration of chromophores (or optical absorbers) due to attenuated light strikes to the superficial layer of tissue. Non-invasive imaging and monitoring in biomedicine based on the developments in photonics for last three decades [11]. The imaging of subsurface of a skin emerges as a very important modality in terms of detection of the many optical properties of skin. Detection of skin cancer under in-vivo conditions based the investigation of non-invasive optical modalities has a capability that further increases the success of sensitivity and image magnification of visible optical window based methods [11]. However, several methods such as Optical Coherence Tomography - OCT, Confocal Scanning Light Microscopy - CSLM, and Magnetic Resonance Imaging – MRI are very expensive and are bulky systems. For this reason, in addition to the studying on visible optical region, at the same time, scanning near-infrared region as well, the developments of cheap and portable systems is becoming a more popular approach.

The aim of this study is to determine the estimation of hemoglobin concentration and oxygen saturation of tissue by non-invasively functional laser imaging for early skin

Received on: 14.03.2017
Accepted on: 21.03.2017

cancer diagnosis. The monitoring of absorption and scattering coefficients that are strongly related to the concentration of chromophores in a skin with light-skin interaction based models is very crucial and is still an open-ended problem.

2. Materials and Methods

2.1. Theoretical Approach

Oxygen saturation is related to the heart rate, the breathing rate, the blood pressure, and the body temperature. It is defined that as relative measure of oxygen amount of dissolved or carried in human body medium. It is important to determine whether a person has adequately supply oxygen. Also the continuous montioring of oxygen saturation is important for detecting hypoxemia condition [12]. For human body, estimation of arterial oxygen saturation at peripheral capillary is called SpO_2 which is primary focus of clinical conditions. SpO_2 is the percentage of oxygenated hemoglobin and expressed as follows,

$$SpO_2 = \frac{HbO_2}{HbO_2 + Hb} x100\% \tag{1}$$

where HbO_2 is the concentration of oxygenated hemoglobin and Hb is the concentration of deoxygenated hemoglobin [12]. Generally, in clincial conditions SpO_2 is measured depending on the selected two wavelengths; λ_1 and λ_2 such that absorbance by HbO_2 is more at at λ_2 than at λ_1 while the absorbance by Hb is more at λ_1 than at λ_2, as given in Figure 1.

Absorption of HbO_2 and Hb molecules depends on wavelength as seen in Figure 1 [13]. Absorption of Hb decreases in contrast to that absorption of HbO_2 increases, as wavelength increases in the range of between 680 nm to 800 nm. The determination of actual absorption coefficient is difficult due to the diffusive nature of light in tissue. A precise quantitative result is complicated by scattering in tissue and absorption of other chromophores in tissue like melanin.

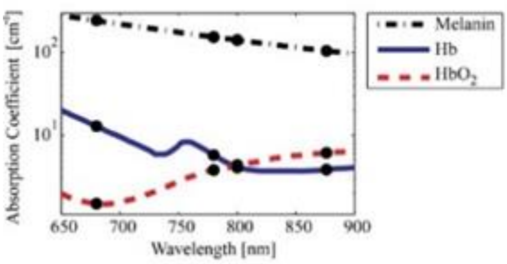


Figure 1. Changes in oxygenated hemoglobin and hemoglobin absorption coefficients according to wavelength

The chromophore distribution effects on backscattered light, so that there is a need a model of light transport in tissue to describe it. The diffusion equation for the absorption and scattering of light is described below for position and time t by [13]:

$$\frac{\partial\phi(\vec{r},t)}{c\partial t} + \mu_a \partial\phi(\vec{r},t) - \nabla\left[\frac{\nabla\partial\phi(\vec{r},t)}{3(\mu_a + \mu_s(1-\mathrm{g}))}\right] = S(\vec{r},t) \tag{2}$$

In the equation above; Φ is fluence rate, μ_a and μ_s are absorption and scattering coefficients, g is the anisotropy factor, and S is an isotropic light source. In the diffusion equation that is given Eq. 2, light is generally attenuated exponentially in a medium for a given depth l, with the wavelength dependent absorption and scattering coefficients as $\mu_a(\lambda)$ and $\mu_s(\lambda)$. In addition, to simplify the problem of estimating absorption, μs is assumed constant on the related spectral bandwidth. If the reflectance data was acquired by a camera, the gray level value (intensity) of one imaged pixel can be described by the value of $A(\lambda)$. The image which is obtained represents the total spatial (x, y) map of backscattered diffuse reflectance.

In our study, red and blue food dyes were used to mimic the oxygenated hemoglobin and hemoglobin chromophores respectively. Considering the variation of the absorbance coefficients of the oxygenated and hemoglobin in relation to the wavelength, it is seen that there is an important difference for the two kinds of hemoglobin at 680 nm as shown in Figure 1 [13].

2.2. Instrumental Setup

The illumination was done by a laser diode which had 680 nm center wavelength and 100 mW power, and the laser diode current was set by the driver circuit to prevent saturation. 1024x768 CCD camera (D223C; Thorlabs, USA) used as a detector. The camera exposure time was set approximately 8 ms and a glass diffuser (DG 100x100-600, Thorlabs, USA) was used in front of the laser to provide a homogeneous surface illumination. The work was performed in the dark room. The instrumental setup is given in the Figure 2.

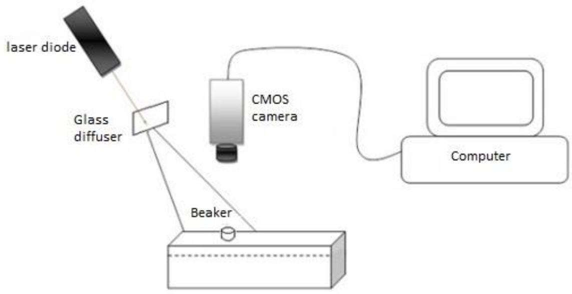


Figure 2. Instrumental setup

2.3. Measurements

The flowchart of the study steps is given and explained in the Figure 3.

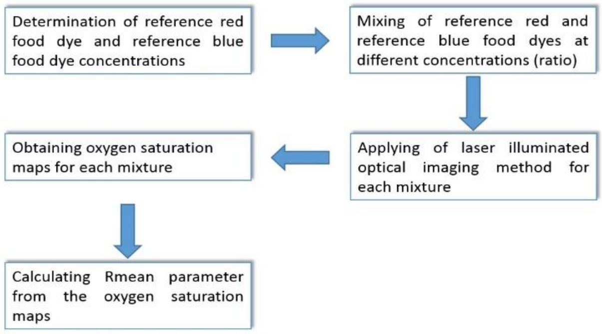


Figure 3. Changes in oxygenated hemoglobin and hemoglobin absorption coefficients according to wavelength

2.3.1. Determination of Reference Red and Reference Blue Dye Concentrations

As a reference material in the study, red food dye (Ponceau4R) and blue food dye (Brillant blue) were used for mimicking oxygenated hemoglobin and hemoglobin, respectively. Red and blue food dyes were put into two different beakers in varying amounts by trial and error. After then, it was mixed with distilled water. The solutions were prepared and poured into two separate regions on glass coverslips and imaged with the camera under ambient light.

Regions of red dye solution and blue dye solution were cropped separately and analyzed by RGB analysis, comparing the "R" and "B" components; histograms of the "R" and "B" components of the red and blue reference solutions in the two beakers were determined to be used in the study; 100 mg/dl and 5 mg/dl respectively. The actual dye mixtures are shown in Figure 4 as an example.

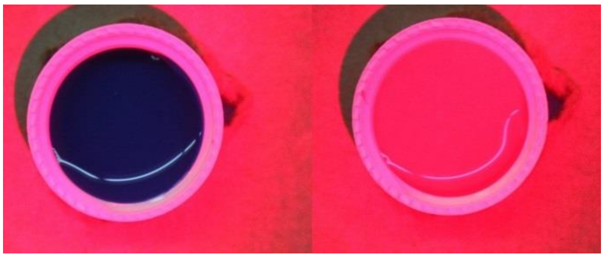


Figure 4. Reference red and reference blue concentrations images

2.3.2. Laser Illuminated Wide Field Imaging

The observation parameter is the average value of the pixel brightness and is essentially equal to the light absorption. Images were acquired from the phantom surface roughly. A rectangular sub-region was cropped for analysis. From the corresponding region image, the R maps (images) were obtained by estimating the "R" parameter corresponding to the oxygen saturation based on the pixel values. For a single pixel, this parameter can be calculated as [4]

$$R = \frac{I_1}{I_1 + I_2} \qquad (3)$$

where I_1 refers to the pixel brightness in the reference red dye at 680 nm and I_2 refers to the pixel brightness in the reference blue dye at 680 nm. The R parameter was generated on a pixel-by-pixel basis for an entire image, thus the saturation map was generated from the "R" value of all pixels.

3. Results

Before the measurement of mixture of the dyes, the red and the blue dye concentraions were considered seperately. From this measurement, the output of the signal was obtained as pixel intensity values averages. The related results are given in Figure 5 and Figure 6 respectively. It is observed that there are inverse relationship between measured optical signal and concentrations.

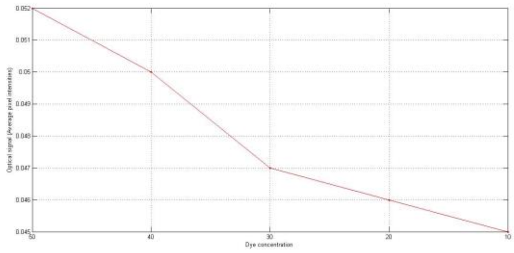


Figure 5. The investigation of the red dye

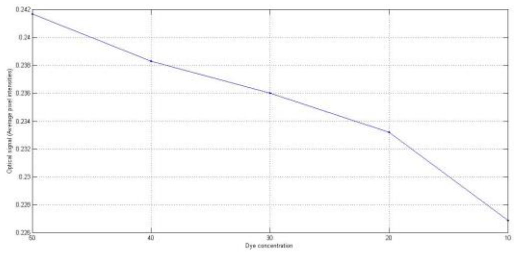


Figure 6. The investigation of the blue dye

The relative concentrations of HbO_2 and Hb were simulated by the dissolution of red and blue dyes in a certain amounts in distilled water respectively. Oxygen saturation (R) maps, produced for different concentrations are given below in Figure 7.

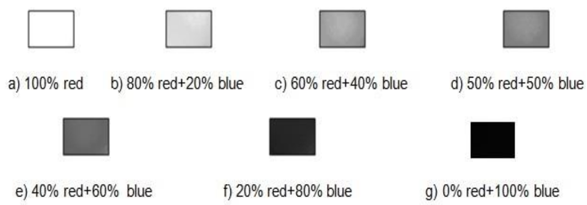


Figure 7. The R maps of the cropped region for different mixtures

In order to enhance the visual observation, the above R maps were subjected to pseudo – coloring which is given in Figure 8.

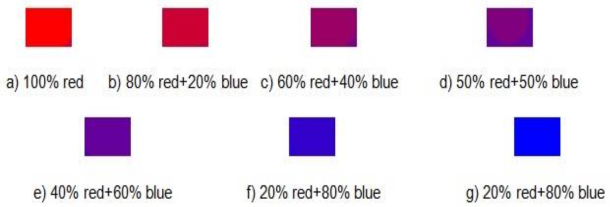


Figure 8. Pseudo – colored R maps

A representative parameter; R_{mean} is generated with calculating an average value from the estimates of oxygen saturation maps (images) which are mentioned above. The variation of the R_{mean} parameter is given in Figure 9. In the following figure, the horizontal axis represents the dye concentrations that corresponding to mixing percentage of HbO and HbO_2.

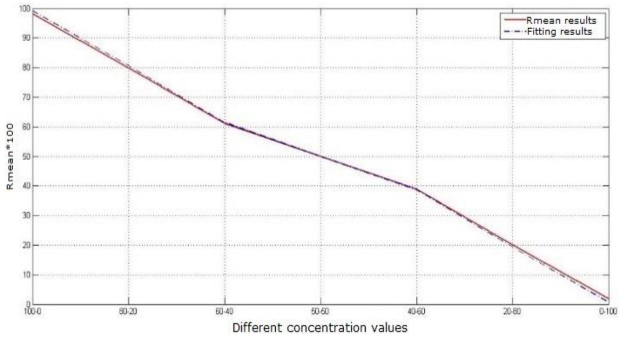


Figure 9. Changing of R_{mean} relating to mixture concentrations

As you can see from the graph; linear variation of the R_{mean} parameter depending on the combination of mixtures is observed. As stated in the literature, the use of 680 nm wavelength is suitable for this situation.

For the use of the behavior obtained for the artificial HbO_2 and Hb concentrations known above in the in-vivo and/or ex-vivo environments, calibration or classification learning stages will be provided by the numerical values produced by the above applications. For this purpose, variations were fitted to a linear equation and RMS fitting error for this equation is around 0.5137.

$$y = 1.0232x - 1.1606 \tag{4}$$

When the obtained results were investigated; it was observed that there was an approximated linear relationship between measured optical signal and dye concentrations. In the litrature like as declared in [6], the same relationship also was observed.

4. Conclusion

In this study, we investigated in-vitro based oxygen saturation measurements where the hemoglobin concentrations were simulated by different red and blue food dye mixing in the distilled water. In clinical conditions, the human tissue or arterial oxygen saturation is measured by two different optical wavelengths method and also photoacoustic technique. However, these methods cause hardware cost and have optical design limitations. Therefore, our proposed single wavelength based study would be promised method.

5. Acknowledge

This study was supported by the “The Scientific and Technological Research Council of Turkey (TUBITAK)” grant project (Tubitak-COST: 113E771).

6. References

[1] Faber D. J. et al. “Oxygen saturation-dependent absorption and scattering of blood”, *Physical Review Letters* 93, 028102, 2004.

[2] van Assendelft O.W., “Spectrophotometry of Hemoglobin Derivatives”, *Royal Vangorcum Assen*, Springfield, 1970.

[3] Zijlstra W.G., Buursma A., and Meeuwsen-van der Roest W. P., *Clin. Chem*. 37, 1633, 1991.

[4] Wray S. et al., *Biochim. Biophys*. Acta 933, 184, 1988.

[5] Tabulated data from various sources compiled by S. Prahl, 30/08/2016, Avaible: http://omlc.ogi.edu/spectra,.

[6] Gao F., Peng Q. W., Feng X. H., Gao B., and Zheng Y. J., “Single-Wavelength Blood Oxygen Saturation Sensing With Combined Optical Absorption and Scattering”, *IEEE Sensors Journal*, vol. 16, no. 7, pp. 1943-1948, Apr 1, 2016.

[7] Kiekkas P. et al., “Knowledge of pulse oximetry: Comparison among intensive care, anesthesiology and emergency nurses”, *J. Clin. Nursing*, vol. 22, nos. 5–6, pp. 828–837, Mar. 2013.

[8] Humphreys K., Ward T., and Markham C., “Noncontact simultaneous dual wavelength photoplethysmography: A further step toward noncontact pulse oximetry”, *Rev. Sci. Instrum.*, vol. 78, no. 4, p. 044304, Apr. 2007.

[9] Millikan G. A., “The oximeter, an instrument for measuring continuously the oxygen saturation of arterial blood in man”, *Rev. Sci. Instrum.*, vol. 13, no. 10, pp. 434–444, Oct. 1942.

[10] Hernández S. E., Rodríguez V. D., Pérez J., Martín F. A., Castellano M. A., and Gonzalez-Mora J. L., “Diffuse reflectance spectroscopy characterization of hemoglobin and intralipid solutions: in vitro measurements with continuous variation of absorption and scattering”, *J. Biomed. Opt.*14 (3), 034026, 2009.

[11] Memorandum of Understanding for the implementation of a European Concerted Research Action designated as COST Action BM1205: European Network for Skin Cancer Detection using Laser Imaging, 2012.

[12]Shao D., Liu C., Tsow F., Yang Y., Du Z., Iriya R., Yu H., Tao N., “Noncontact Monitoring of Blood Oxygen Saturaiton Using Camera and Dual-Wavelength Imaging Ssytem”, *IEEE Trans. On Biomed. Eng.*, 63 (6), 1091-1098, 2016.

[13] Dhawan A. P., D'Alessandro B., “Multispectral Transillumination Imaging of Skin Lesions for Oxygenated and Deoxygenated Hemoglobin Measurement”, *New Jersey Institute of Technology*, Newark, NJ, 2010.

A New CPW-Fed Circularly Polarized Square Slot Antenna Design

Saeid KARAMZADEH[1,2], Hemrah HIVEHCHI[1]

[1]Department of Electrical and Electronics Engineering, Istanbul Aydin University, Istanbul, Turkey
[2]Application & Research Center for Advanced Studies, Istanbul Aydin University, Turkey
karamzadeh@itu.edu.tr, hemrahhivehchi@stu.aydin.edu.tr

Abstract: *In this work, a new wideband circularly polarized square slot antenna (CPSSA) with a coplanar waveguide (CPW) feed is proposed. The suggested antenna is wrought of one arc-shaped, two inverted-L grounded strips around, two opposite corners of the square slot and C shaped(smile) gap on the patch. In this outline the impedance bandwidth and the axial ratio bandwidth (ARBW) are increased compared to the previous CPSSA designs. In this way, the 3 dB axial ratio bandwidth of the designed antenna improved and reached to more than 57 % (3.2 GHz - 5.7 GHz) and the VSWR (Voltage Standing Wave Ratio) < 2 impedance bandwidth increased as large as 108 % (3 GHz - 10.1 GHz). Overall this work, the optimized method of the axial ratio (AR) and S11are provided and discussed in text. The fabricated antenna results agree well with the simulation analysis.*

Keywords: *CPSSA, CPW and C shaped gap.*

1. Introduction

A microstrip slot antenna is a good offer for communication systems, radar and remote sensing applications, as it is low specifications, low cost, lightweight, and can be easily integrated with monolithic microwave integrated circuits (MMICs) [5]. To prosper the operating bandwidth and not decrease the antenna size, applying the printed slot antenna is a feasible method. In this technology circularly polarized (CP) has been suggested more than liner polarized (LP) antenna [5]. CP antennas are more attention than other importance systems in wireless communications, sensors, radio frequency identifier (RFID), and vehicular radar. CP antennas is a good choice among the various designs and structures in wireless communications to raise of system implementation offering better mobility, and weather penetration, more than the linearly polarized (LP) antennas [1, 10,13,14,15 and 16]. Also by applying the CP antennas in wireless communication systems, arranging the direction of the antenna between the transmitter and receiver is not needed any more [6]. CP antennas overcome the multipath fading problem and enhance system performance. Through feeding methods, coplanar waveguide (CPW) feed has some advantages like wide bandwidth, easy integration and single metallic layer [7]. In order to producing CP radiation, created some of the methods such as : implementing two inverted-L grounded strips around two opposite corners of the slot [1,8], improvising T-shaped grounded metallic strip, which is orthogonal to the axis of the CPW feed-line [9], using an asymmetrical CPW fed from a corner of the slot with an embedding pair of grounded strips implanted in the slot [12], striped slot antenna with longer fraught [10], utilizing the additional arc-shaped grounded metallic strip for circular and linear polarization [11], and embedding a firelight shaped feed line and inverted-L grounded strips [3,8] are used in literature. In this work a new design of a CPW fed circularly polarized square slot antenna by combining previous methods and improve some part for getting better results is presented. Based on simulated results, the impedance bandwidth is about 108 %, totally inclusive the 3 dB AR bandwidth, which is about 57 %. This antenna is appropriate for IEEE 802.11a, (5.15–5.35GHz / 5.72–5.82 GHz) and for IEEE 802.16, (3.2–3.8 GHz / 5.2–5.8 GHz).

2. Antenna Design

The configuration of provided antenna has shown in figure 1. This antenna is printed on a commercially cheap FR4-epoxy substrate with $\varepsilon r = 4.4$, $\tan(\delta) = 0.024$. The feed line of the proposed antenna is CPW and is connected to a 50 Ω SMA connector. The gap between the feed line and the ground is 0.3 mm, which is widened at the end to improve the impedance bandwidth. The CP operation of the proposed antenna is greatly related to the one arc shaped and two inverted-L grounded strips in opposite sides placed around the corners of the square slot with C shaped gap on the patch. The antenna parameters are as follow respectively: G = 60, L = 40, h = 0.8, L1=10, L2 = 12, R1 = 11, R2 = 9, L3 = 1, W3 = 4.6, L4 = 1, W4 = 18.4, L5 = 11.6, L6 = 9.8, L7 = 9.6, L8 = 7.9, R3 = 7.5, W1 = 4, W2 = 3.15, L9 = 13, L10 = 6.7, L11 = 6, and W5 = 3.6 (All units are in millimeters).

Received on: 23.05.2016
Accepted on: 04.08.2016

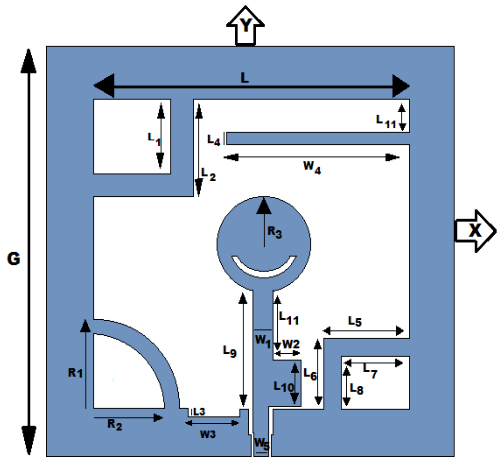


Figure 1. Configuration of the proposed antenna.

3. Simulation and Measured Results

In this section we provided antenna in six prototypes with simulations results for every prototype and measured for the final version. For designing antennas step by steps and tracing the results in each footstep Ansoft high frequency structure simulator software (HFSS, ver.16) has been used. For simplification in the antenna design G = 60 mm, L = 40 mm, h = 0.8 mm were already selected. Figure 2 clarifying the antenna designing process. Ant. 1 includes only a feed line connected to semicircular patch and ground plane; Ant. 2 includes two inverted-L grounded strips around top and bottom corners and one arc-shaped grounded strip at the bottom left side corner. In Ant. 3 a vertical tuning stub (L10 × W2) has been located in the feed structure and the gap between the signal strip and the ground plane is widened at the end. Ant. 4 has a tuning slit (L3 × W3) that has been removed from the ground plane at the left side of the feed line. In Ant. 5 a horizontal grounded strip (L4 × W4) embedded at top right corner and in Ant.6 C shaped gap created on the radiation patch.

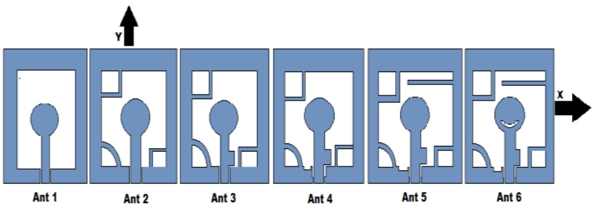


Figure 2: Antenna designing process.

Figure 3 and figure 4 indicate the simulated results for frequency responses of -10 dB S11 and 3 dB axial ratio variations for the six provided prototypes respectively. From Fig. 4, it could be seen that Ant. 1 has a linear polarization. By embedding two inverted-L and an arc-shaped grounded strip around the corners in ant 2, the AR is amended, in Ant 3, the AR is greatly better which reaches 5 % (4.8 GHz–5.4 GHz). However, in this case the AR is not guaranteed by -10 dB S11. A vertical tuning stub is added to the feed structure and the gap of CPW feed is organized to a step shape for improve the impedance matching.

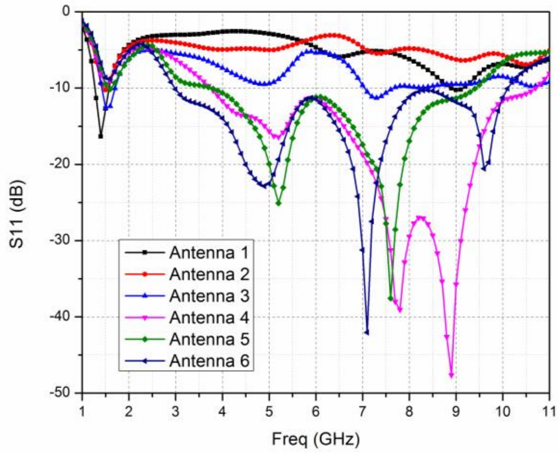


Figure 3: Simulated results of S11 for the six provided prototypes.

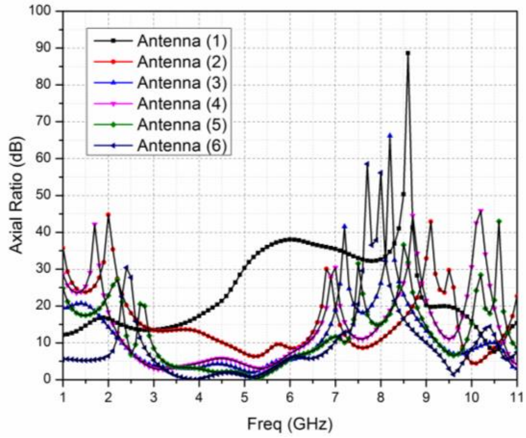
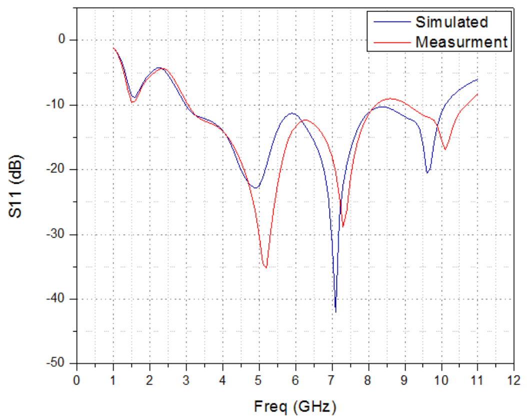


Figure 4: Simulated results of AR for the six provided prototypes.

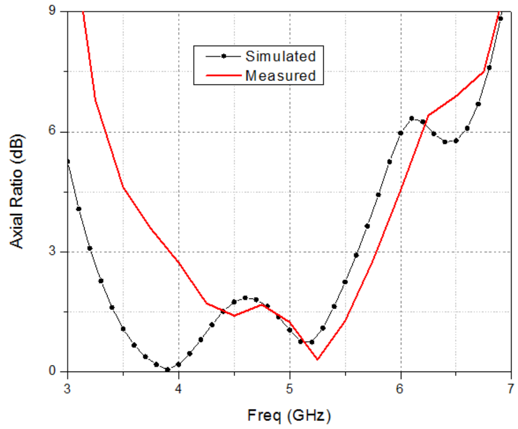
As shown in Fig 4, these structures have great effect on the impedance bandwidth of Ant. 2 (called Ant. 3). With embedded a rectangular slit the AR will be decreases 0.15% (3.017- 3.081) in Ant.4. In Ant. 5 by adding a horizontal strip at the top right corner after tested different length and choice a best length for antenna optimization of the ground plane the AR bandwidth will reach about 20 % (3.8 GHz – 5.7 GHz). At last, by embedding the semicircular gap at the patch of the feed, not only AR bandwidth is increased to 57 % (3.2 GHz – 5.7 GHz) but also the impedance bandwidth can be increased to cover the whole UWB bandwidth. The simulated results in Fig. 3 indicate that including the horizontal strip and C shaped gap, greatly influence the ARBW.

Figure 5 (a) and (b) illustrate the simulated S11 and AR characteristics for the designed antenna at least prototype. Close communication between the simulated and measured results is observed. As also indicated in Fig. 5 (a), the measured impedance bandwidth of the proposed antenna is from 3 GHz up to 10.1 GHz (108 %) for VSWR < 2 and the measured 3 dB ARBW is increased from 3.2 GHz to 5.7 GHz (57 %) that is about 2.5 GHz. Due to the AR and the impedance bandwidth.

In Table 1 the impedance bandwidth, AR bandwidth of the designed antenna has been compared with related work in literature [1, 2, 3]. It is observed that the proposed antenna has wider ARBW and impedance bandwidth than the other ones. All these antennas were fabricated on an FR4 substrate with a loss tangent of tan (δ) = 0.024, permittivity of $\varepsilon_r = 4.4$.



(a)



(b)

Figure 5: Measured and simulated diagrams of (a) S11 and (b) the AR of the proposed antenna (Ant 6).

Table 1. Compared of the measured characteristics of some CPSS antennas with proposed work

References	F_c (GHz)	Imp.Band (GHz)	ARBW (GHz),%
Ref.1	2.665	1.6-3.055	(2.3-3.03), 27.4%
Ref.2	5.969	2.674-13.124	(4.9-6.9), 32.2%
Ref.3	2.754	2.023-3.424	(2.07-3.4), 48.8%
Proposed Work	7.1	3-10.1	(3.2-5.7), 57.47%

The simulated and measured gain of proposed antenna is shown in figure 8.

Simulated radiation pattern for different frequencies and phases has been presented in figure 9 also.

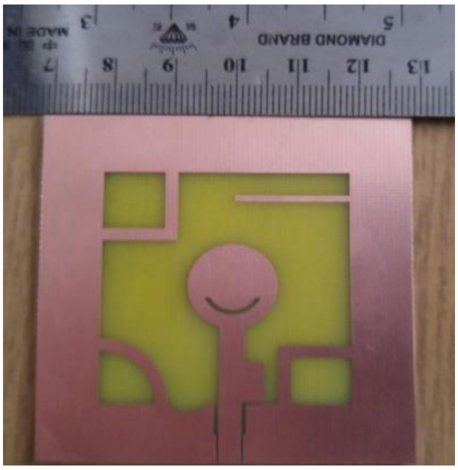


Figure 6: Photograph of the realized CPSS antenna.

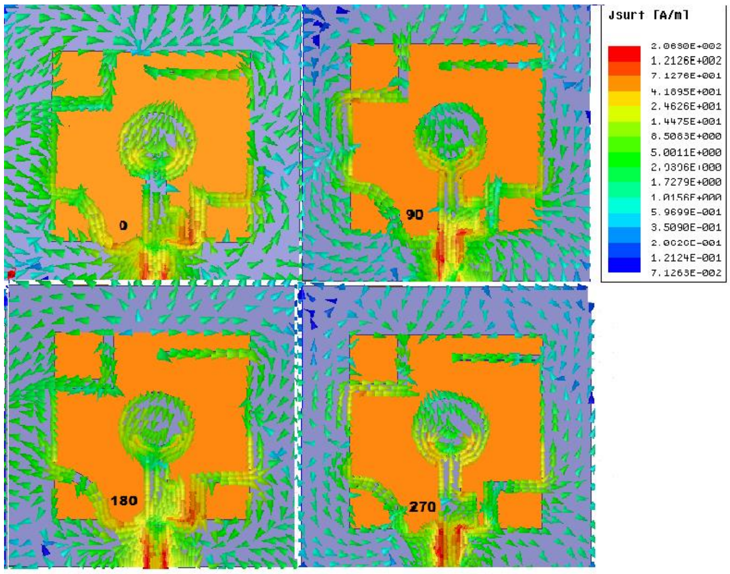


Figure7: Current distribution of the proposed antenna at 6.5 GHz in 0, 90, 180 and 270 phases.

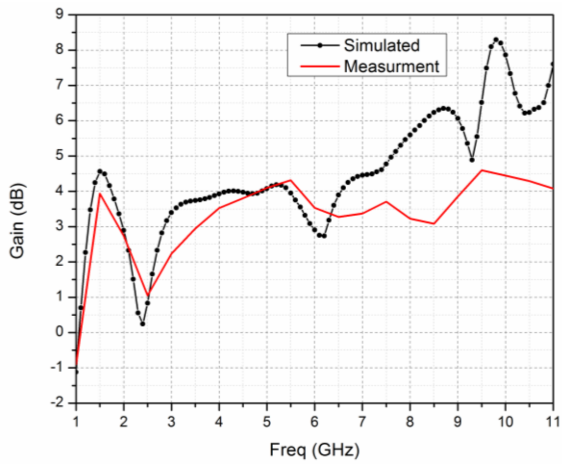
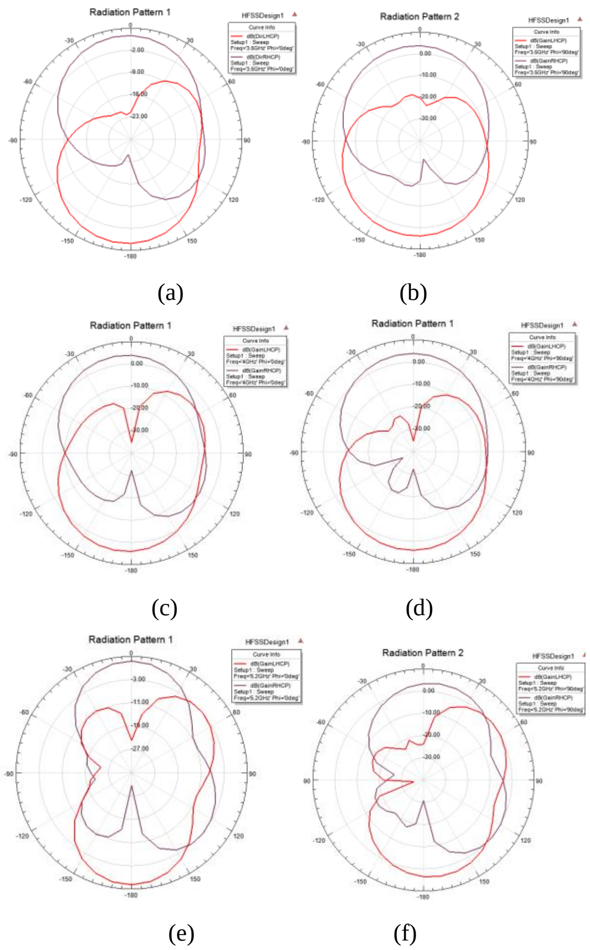


Figure 8: Measured and simulated antenna gain of proposed antenna.



(a) (b) (c) (d) (e) (f)

Figure 9: Simulated radiation patterns of the CPSS proposed antenna at:(a) 3.5 GHz, 0^0, (b) 3.5 GHz, 90^0, (c) 4 GHz, 0^0, (d) 4 GHz,90^0, (e) 5.2 GHz, 0^0, (f) 5.2 GHz,90.

4. Conclusions

This paper proposed a new wideband CPW-fed CPSSA for Ultra Wide Band (UWB) applications. In the antenna geometry, employing one arc-shaped, two inverted-L grounded strips and a C shaped gap on the patch, improved 3 dB ARBW and impedance bandwidth noticeably in compare with previous works. The obtained results show that the proposed antenna has an impedance bandwidth about 7 GHz (108 %) and a 3 dB AR bandwidth of about 2.5 GHz (57 %).

5. References

[1] J. Y. Sze and C. C. Chang, “Circularly polarized square slot antenna with a pair of inverted-L grounded strips,” *IEEE Antennas Wireless Propag. Lett*., vol. 7, pp. 149-151, 2008.

[2] J. Pourahmadazar, C. Ghobadi, J. Nourinia, N. Felegari, and H. Shirzad., “Broadband CPW-Fedcircularly polarized square slot antenna with inverted-L strips for UWB applications,” *IEEE Antennas Wireless Propag. Lett*., vol. 10, pp. 369- 372, 2011.

[3] J. Y. Sze, C. I. G. Hsu, Z. W. Chen, and C. C. Chang, “Broad band CPW-fed circularly polarized square slot antenna with lightening-shaped feedline and inverted-L grounded strips,” *IEEE Trans.Antennas Propag*., vol. 58, no. 3, pp. 973–977, Mar. 2010.

[4] B. Boroomandisorkhabi, R. A. Sadeghzadeh, F. B. Zarrabi, and E. Ghahramani “A Novel UWB Circular CPW Antenna with TripleNotch Band Characteristics” *IEEE Trans Antennas Propag*, pp. 978, Nov. 2013.

[5] D. S. Javan and O. H. Ghouchani, “Cross slot antenna with u-shaped tuning stub for ultra wideband applications,” *Appl. Comp. Electro.Society (ACES) Journal*, vol. 24, no. 4, pp. 427- 432, Aug. 2009.

[6] K. L. Wong, F. S. Chang, and T. W. Chiou, “Low-cost broadband circularly polarized probe-fed patch antenna for WLAN base station,” I*EEE Antennas Propag. Soc. Int. Symp.*, 2002, vol. 2, pp.526–529.

[7] C. C. Chou, K. H. Lin, and H. L. Su, “Broadband circularly polarized cross-patch-loaded square slot antenna,” *Electron. Lett.*, vol. 43, no. 9,pp. 485–486, Apr. 26, 2007.

[8] Jia-Yi Sze, Chi-Chaan Chang “Circularly Polarized Square Slot Antenna With a Pair of Inverted-L Grounded Strips” *IEEE Antennas And Wireless Propagation Letters*, VOL. 7, 2008.

[9] Saeid Karamzadeh, Vahid Rafii, Mesut Kartal, and Morteza Dibayi “Circularly Polarized Square Slot Antenna Using Crooked T-Shape Technique” *Applied Computational Electromagnetics Society Journal (ACES)*, Vol. 30, No. 3, March, 2015.

[10] N. M. Jizat, S.K.A. Rahim, Y.C.Lo, M. M. Mansor “Compact Size of CPW Dual-Band Meander-Line Transparent Antenna for WLAN Applications” *IEEE*, Dec 2014

[11] M. J. Chiang, T. F. Hung, and S. S. Bor., “Dual band circular slot antenna design for circularly and linearly polarized operations,” *Microw. Opt.Technol. Lett*., vol. 52, no. 12, pp. 2717-2721, Dec, 2010.

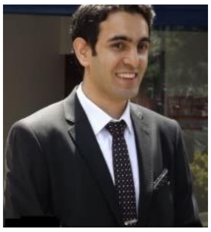
[12] J. Y. Sze, J. C. Wang, and C. C. Chang, “Axialratiobandwidth enhancement of asymmetric CPW fed circularly-polarized square slot antenna,” *Electronics Lett*., vol. 44, pp. 1048-1049, 2008.

[13] S. Karamzadeh, Bal Virdee, V. Rafii, M. Kartal, “Circularly polarized slot antenna array with sequentially rotated feed network for broadband application ”, *International Journal of RF and Microwave Computer-Aided Engineering*, vol. 25, no.4 pp. 358-363, May, 2015.

[14] S. Karamzadeh, M Kartal “Circularly polarised MIMO tapered slot antenna array for C-band application”, *Electronics Letters*, Vol.51, no.18, pp. 1394 -1396, August, 2015.

[15] S. Karamzadeh, M. Kartal, “Circularly polarized 1× 4 square slot array antenna by utilizing compacted modified butler matrix and branchline coupler ” *International Journal of RF and Microwave Computer-Aided Engineering*, Vol.26, no:2, pages 146–153,October, 2015.

[16] S Karamzadeh, V Rafii, M Kartal, H Saygin, “Compact UWB CP square slot antenna with two corners connected by a stripline ”, *Electronics Letters*,Vol.52, no.1, pp.10-12, December, 2015.



Saeid Karamzadeh received his MS and Ph.D. degree in Department of Communication Systems, Satellite Communication & Remote Sensing program at Istanbul Technical University in 2013 and 2015 respectively. He won the award of the most successful PHD thesis form Istanbul Technical University. He is an Assist. Prof. Dr. in the Istanbul Aydin University, Department of Electric and Electronics Engineering and also with Application & Research Center for Advanced Studies, Istanbul Aydin University, Turkey. His research interests include remote sensing, radar, microwave and Antenna design.

Hemrah Hivehchi was born in Gonbade Kavoos, Iran in 1984. He received his B.Sc. degree in Electrical and Electronic Engineering from Azad University, Aliabad katool, Iran, in 2011 and he is M.Sc. student in Electrical and Electronic Engineering in Istanbul Aydin University now. His research intrest is about microstrip antenna design.

A NOVEL REVERSIBLE FAULT TOLERANT MICROPROCESSOR DESIGN IN AMS 0.35UM PROCESS

M. Hüsrev CILASUN[1], Mustafa ALTUN [2]

[1]Aselsan A.Ş., Gölbaşı, Ankara, Turkey
mhcilasun@aselsan.com.tr
[2]Istanbul Technical University, Maslak, Istanbul, Turkey
altunmus@itu.edu.tr

Abstract: *In this study, reversible circuits are revisited to achieve extreme soft-defect awareness in classical CMOS circuits. Defect models in the literature are reviewed and defect scattering is analyzed. A reversible 8-bit full adder is designed in 12-bit block code domain. As a proof of concept, a pair of reversible ALUs are embedded into a microprocessor with block-code encoded data-path. The design is simulated in ams 0.35um process and a layout is obtained for tapeout.*

Keywords: *Reversible Computing, Fault Tolerance, Microprocessor.*

1. Introduction

Fault tolerance is an important concept for critical circuits operating in harsh conditions with high reliability demands. When a bit error is occurred in runtime, it is very hard to detect and fix the abnormal circuit behavior. In most cases, soft errors happen instantaneously and it is unlikely that a permanent fault pattern is observed. Soft error concept has been known since early 1970s and it might be caused by various phenomena.

In very early years of satellites, several errors in circuits are observed [1]. However, these errors were not related to charge accumulation in capacitors due to solar winds. The errors were found to be as a consequence of high energy particles in deep space. Although the initial attitude is that the terrestrial circuits are safe against these heavy ions which cannot survive in the world's atmosphere, further research has shown that these ions can trigger a reaction chain which eventually produces failures in sea-level electronic circuits [2]. Although the error rate of one error in several years might sound trivial, such a failure might be extremely critical for military or space applications. Alongside these, radioactive impurities in either packaging or doping material can also result in soft errors [3].

Alpha particle emission from the impurities in the packaging material, high-energy protons and neutrons triggered by cosmic rays, thermal neutrons, random background noise, and signal integrity (SI) problems might cause soft errors. Soft errors, such as SEU (Single Event Upset) and SET (Single Event Transient) and their correction has critical significance in terms of space environment considerations. In this sort of applications, the circuit is designed using DMR (Dual modular redundancy) and TMR (Triple modular redundancy) that is widely utilized as an irreversible design methodology, yet these approaches, combined with the inherent fault tolerance of traditional CMOS logic design, makes it harder to track and detect error patterns throughout the signal. In the literature, many circuit synthesis algorithms and test methodologies are developed and several realization techniques are suggested. In this paper, these areas will be combined, by taking CMOS reversible circuit realization as a base approach, a new circuit synthesis method is developed. Using this method, an 8-bit microprocessor is designed by considering previously suggested test methods. The design will include a fault tolerance model by taking advantage of block code mapping which will lead to the detection of multiple errors. Thus, the processor will be available for on-line testing. After optimizing the design and obtaining the realization of the circuit, testing process will begin.

To keep the paper self-containing, several preliminary information on reversible circuits will be presented. A reversible circuit is traditionally defined as a bijective mapping between two identical *n*-dimensional Boolean spaces. As their name suggest, the very basic distinctive property of reversible circuits is the availability of a backward mapping for any possible input combination in contrast with conventional combinatorial logic circuits. As a property inherited from quantum circuit design, reversible circuits are mostly considered as combinatorial cascades of reversible gates. Reversible circuits and gates can be represented in various ways. The most common form is the reversible circuit diagram which shows the cascades of reversible gates that are applied in corresponding lines. However, equivalent gate and circuit functionality can be described as a permutation matrix, a decision diagram as well as a truth table. If the gate diagram is used, an *n*-line logic vector can be propagated from the input to the output by

Received on: 28.05.2016
Accepted on: 02.11.2016

applying the function of each corresponding gate. If the permutation matrix representation is considered, the overall circuit functionality can be obtained as follows: if the gates are in series, a cross-product of gate matrices is considered. If the gates are in parallel, a tensor (Kronecker) product of gate matrices is computed.

The very first universal reversible gate is a CCNOT or a Toffoli gate which is proposed by Tomasso Toffoli in 1980 [4]. Toffoli gate has three inputs and three outputs. It basically flips the third input if the first two inputs are both one. By modifying the inputs to make its truth table corresponding to AND and NOT gates, one can easily show that the Toffoli gate is universal. Toffoli gates can be generalized by increasing the number of control lines (which are the black dots in the first two lines) as well as the number of targets. In [5], multiple-target Toffoli gates, namely mEXOR gates are introduced for this purpose. Toffoli gates are at the main focus of reversible circuit research since they are convenient for synthesis and easy to implement.

Scientific interest on reversible circuit synthesis dramatically increases with the introduction of quantum computing [6,7,8] since the traditional circuits are believed to be trapped by Moore's limit [9]. Since the reversibility is a must for quantum circuits due to physical obligations, numerous synthesis methods have been offered [5,10,11,12]. In synthesis, the essential concern is minimizing the number of gates, number of lines, and keeping the synthesis time as small as possible. However, the circuit is often described as truth tables whose lengths are proportional to the exponent of the number of lines, which makes the synthesis runtime infeasible for considerably high number of lines [13]. Also, reversible circuits are believed to put a lower - theoretically zero - power limit to logic circuits [14,15,16]. Reversible CMOS circuits are previously realized in pass-transistor logic with no promise on a lower bound in power consumption [17,18].

In their seminal paper in 2003, Miller *et al* proposes a novel idea for reversible circuit synthesis [10]. Their algorithm takes the truth table as an input, and processes row by row. In each row, it is guaranteed that at least one row is matched between the input and the output. Although this method is far from optimal, it is good in terms of the number of lines and avoiding garbage generation. Further transformation based methods have also been developed. For instance, in [10] there are given cyclic equivalency relations of certain cascades of Toffoli gates. A *template*, is defined as a Toffoli network whose function is an identity mapping. When cascades are replaced with the smaller template counterparts, size of the overall circuit is reduced significantly.

Although the transformation based methods are practically useful, they represent the initial data as truth tables which are exponential in size. Very long synthesis durations can be reduced [19], yet different approaches are required for a circuit-level optimality. The size of the data complicates getting a more compressed representation of the same functionality. In [11], reversible circuits are expressed in terms of binary decision diagrams and their synthesis results are better than previously proposed methods in terms of the size and computation time, which makes reversible circuit synthesis scalable especially for large functions. In [12], the decision diagram structure gets complicated, which allows a general set of Quantum circuits to be synthesized in a scalable manner. These diagrams are called QMDD, which stands for Quantum Multiple-valued Decision Diagram. Table 1 summarizes different methods and their time complexities. In designing our microprocessor, we have benefited from these techniques.

Table 1. Worst-case complexities the four possible ways to represent reversible circuits

Method	Complexity
Reversible Circuit Diagram	$O(mn)$
Truth Table	$O(2^n)$
Permutation Matrix	$O(2^{n^2})$
QMDD	$O(n)$

This paper is organized as follows. In Section 2, we give introductory information and make analysis on defects in reversible circuits. In Section 3, we present our microprocessor design. In Section 4, we present simulation results and elaborate on them. In Section 5, we discuss our contributions and future works.

2. Defects and Reversible Circuits

A *defect*, *fault* or *error* is a generic concept which describes undesired behavior of designed model due to internal or external abnormalities. In the content of electrical circuits, a defect is frequently related to unexpected voltage, current, charge or flux characteristics. Fault analysis is a well-studied topic for traditional logic circuits. The main classification of defects is based on the transience of the behavior. If the fault happens only for once and unlikely to reproduce, it is called a *soft error*. Similarly, if the fault causes a permanent change in circuit behavior, the effect is a *hard defect*. One of the most studied types of soft defects is Single Event Upset (SEU) which describes a bit flip in a node of a logic circuit. For a stricter and more detailed classification of soft errors in irreversible circuits, the reader is referred to [3].

From the quantum computing perspective, the error is mostly due to local decoherence or quantum noise. Since redundancy is not allowed [20], quantum computers rely on different error correction schemes in order to operate properly. Although the irreversible circuit faults are classified both in abstract and technology dependent manner, reversible circuit defects are only conceptually classified due to uncertainty or immaturity of reversible implementation schemes.

In a reversible circuit, bit flipping simply refers to the logical inversion of the value of a certain node, among a large variety of fault models. Since bit flipping is generic and can be generalized for modelling other defect models, it will be considered as the base model throughout this paper.

Since witnessing a soft defect in conventional circuits is extremely rare event, conventionally errors are modeled as a single fault in overall circuit. However, especially from the reversible circuit perspective, a single error is not necessarily what is observed at the output. Before proposing a way to detect or correct errors that can happen in reversible circuits, it is essential to investigate the quantitative behavior of a single error. When we inject random single errors to four different circuits, the number of errors at the output varies as the histogram in Figure 1 suggests.

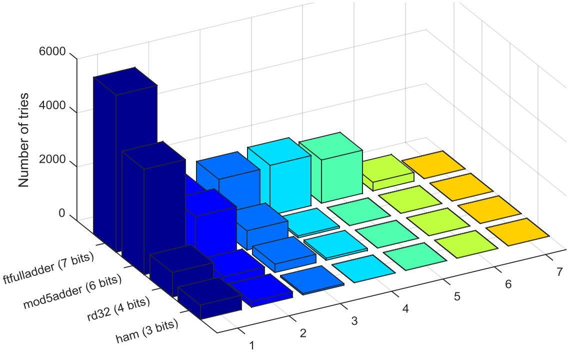


Figure 1 Unnormalized fault scattering histograms of four testbench circuits. Circuits ham, rd32, and mod5adder are taken from [21].

Since the logic applied in reversible circuits is inherently traceable, reversible circuits have a great potential to detect erroneous patterns that occur during circuit operation. In one of the earliest work on reversible fault awareness, Parhami [22] suggests parity preserving reversible gates for detecting erroneous circuit operation. In [22], it is proposed that gates such as Fredkin Gate (FG) and Feynman Double Gate (FRG) can be used for fault awareness since they have parity preserving property which means parity of inputs and outputs are equal to each other for any possible input applied to the circuit. Since each parity preserving gate is considered as a black box, the overall circuit consists of parity preserving gates will also be parity preserving. If there is a single bit flip at any node of the circuit, the corruption in the intermediate parity will propagate to the output, which will eventually cause a mismatch between input and output parities. However, there are several issues on so-called "fault-tolerance" proposed in [22]. Similar to the several papers [23] in the literature, "fault-tolerance" term is loosely defined since it actually implies an "awareness" of the fault pattern, such as described in [24]. It would be a more accurate to classify parity-preserving gates as "defect-aware" or "error-detecting" entities. Alongside this, since the parity-preserving gates are considered as black boxes, it is assumed that only one bit can be erroneous, thus no internal gate fault can result in multiple faulty lines at the output of a parity preserving gate. However, this technology-independent abstract consideration ignores the fact that the newly defined parity-preserving gates are too complex to be just a "black box". Recent parity-preserving gates can be simulated using several Toffoli gates. If an error occurs at the middle of a Toffoli cascade, there is no guarantee that it will not scatter into multiple bit errors at the output of the gate. Therefore, the initial single error assumption collapses.

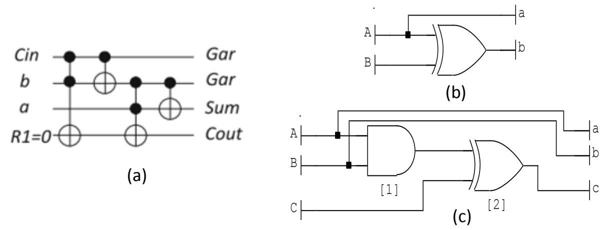


Figure 2 (a) A reversible full adder, adapted from [21]. (b) CMOS implementation of CNOT (Feynman) gate. (c) CMOS implementation of CCNOT(Toffoli) gate.

Consider the reversible circuit in Figure 2(a) and assume that it is implemented using CNOT and CCNOT gates presented in (b) and (c), respectively. If the error is at the inputs of XOR gates, the error will be seen at the output. If the error is at one of the inputs of the AND gate, AND gate might tolerate the error, yet still the fault is propagated to the output. Therefore, reversible CMOS ensures perfect fault awareness while conventional CMOS fails to accomplish this.

3. Microprocessor Design

In order to prove the concept which has been discussed in the previous chapter, a microprocessor is a comprehensive challenge. Since a low-budget fabrication is also planned, the microprocessor should be kept as small as possible. The resulting chip will be sent to *Europractice* for Multiple Project Wafer (MPW) runs which enables relatively cheap prototyping. The designed circuit will utilize a single memory for both instructions and data, therefore it can be considered to have an unpipelined Von Neumann architecture. This is illustrated in Figure 3.

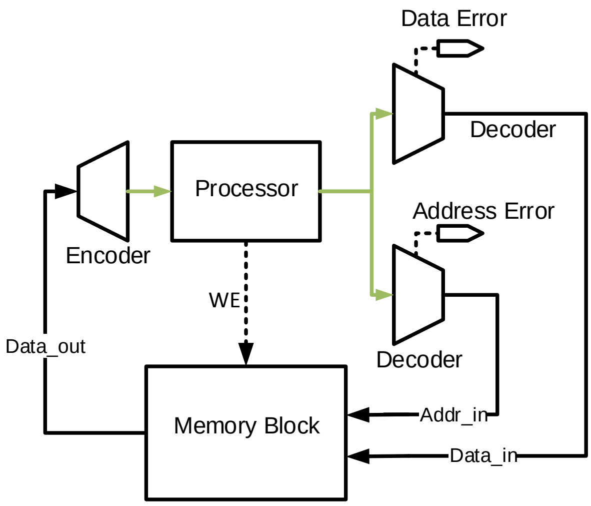


Figure 3 The top operating scheme of the CPU. Clock and reset input pins are not shown. The green path is the way where 8-bit data flow is 12-bit block-code encoded. The dashed signals are single wires which are not included in a bus.

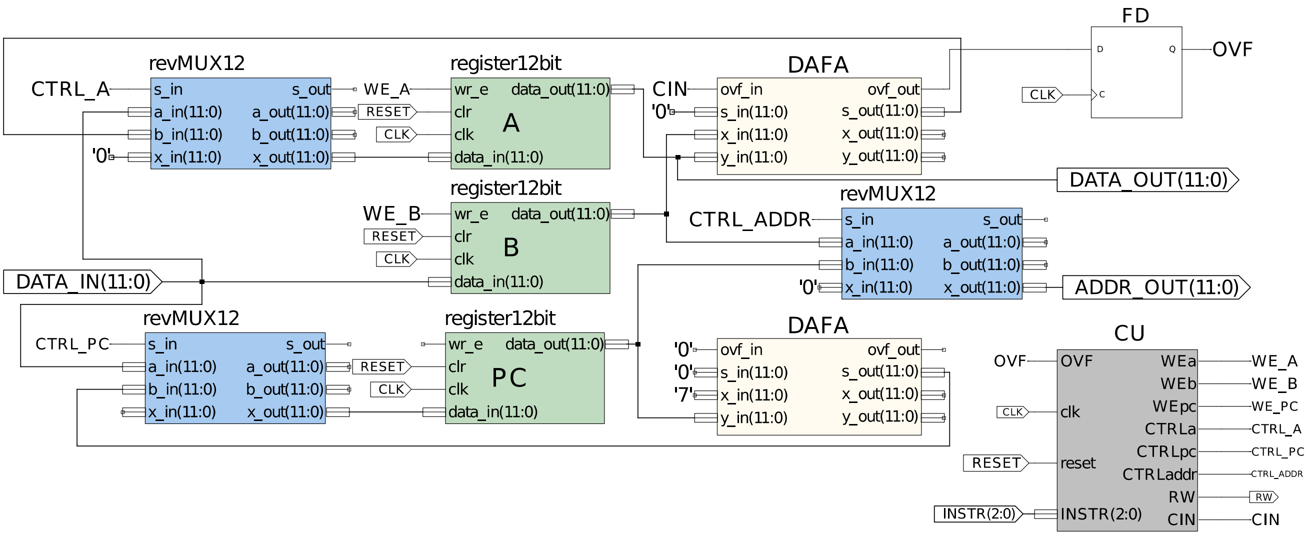


Figure 4 The functional block diagram of the microprocessor.

The instruction set is as small as possible to be operated sufficiently such that the device is classified as a RISC processor. In order to achieve fault-tolerant property, the input will be given as a 12-bit bus instead of actual 8-bit data bus in a block code encoding, as previously proposed for quantum circuits [25]. The accumulator A, register B and program counter register PC will store data and address information as a 12-block code. The ALUs will also be capable of performing reversible addition and subtraction of 12-bit block code encoded 8-bit unsigned integers. Due to the reversibility, no errors at the ALUs will disappear as discussed before. Also, the errors at registers will also be propagated to the output. The functional block diagram of the microprocessor is given in Figure 4.

In the practical operation scheme, a block code encoder and decoder will accompany the microprocessor alongside a Von Neumann styled memory. The top operation scheme is illustrated in Figure 3 where, the encoder performs block code mapping of the 8-bit input data or address onto 12-bit block code domain. When the data are being read from the microprocessor, two decoders perform maximum likelihood decoding to recover the data. Since $d = 3$, decoders will correct one error and also raise an output signal up to two errors, indicating that the output might not be reliable. The memory has a 256 bytes of storage with positive write enable input and positive asynchronous reset. The three most significant bit of the 8-bit raw output of the memory is the instruction input of the microprocessor while the encoded part is connected to the data-in pins.

The design includes three conventional 12-bit multiplexers. In future work, MUXes are planned to be replaced with reversible counterparts in order to ensure maximum defect awareness. Multiplexers are controlled with control signals CTRL_A, CTRL_PC and CTRL_ADDR by the control unit (CU).

When CTRL_A signal is low, the input of the accumulator is connected to the output of the ALU and when CTRL_A is high, input of the accumulator is received from the external data input. If CTRL_PC is low, program counter loads the 1-incremented value of itself from the previous cycle and it loads the external address if otherwise. If CTRL_ADDR is low, the output address bus forwards the content of the program counter PC. If it is low, then content of the register B is forwarded. WE_A, WE_B and WE_PC signals are the Write-Enable inputs of the A, B, and PC registers, respectively. All registers have positive clock and positive asynchronous reset inputs as well. The external read/write signal RW and ALU's input carry signal, CIN are also driven by the control unit. If the subtraction will be performed, CIN is raised, but the second input B is expected to be stored as an inverted manner since the ALU has not controllable inverter. There are two reversible DAFA modules where the first DAFA is the arithmetic logic unit connected to A and B registers. The second DAFA module is connected to the PC register and a constant-1, thus it behaves as a program incrementor.

Table 2. Instruction set of the microprocessor.

Instruction	Operation	Cycles
ADD [000x xxxx]	A <= A+B PC <= PC+1	2
SUBTR [001x xxxx]	A<=A-INV(B) PC <= PC+1	2
LDA [010x xxxx; DATA_IN]	A<=DATA_IN PC <= PC+2	2
LDB [011x xxxx; DATA_IN]	B<=DATA_IN PC <= PC+2	2
STR [100x xxxx; ADDR_IN]:	B<=ADDR_IN PC <= PC+2 Enable W/R	3
JMP [101x xxxx; ADDR_IN]	PC<=ADDR_IN	2
JMPOVF [110x xxxx; ADDR_IN]	PC<=ADDR_IN Overflow==1	2
HALT [111x xxxx; ADDR_IN]	Halts execution.	1

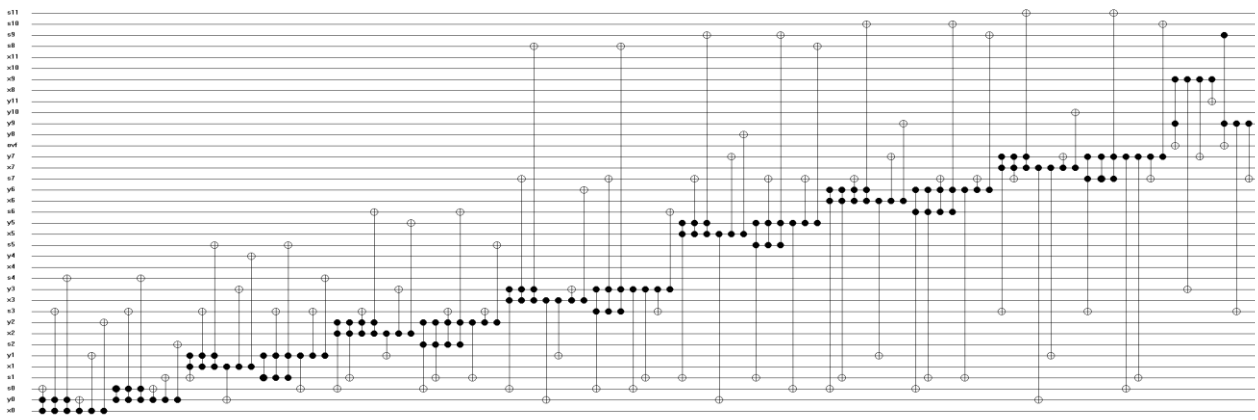


Figure 5 Reversible block code mapped fault tolerant 8-bit full adder based on [26].

The final reversible circuit of the processor composed of Toffoli gates is shown in Figure 5.

The microprocessor has eight instructions as given in Table 2. Although the instruction set is quite limited, it is sufficient to be a proof of concept. The control unit is described in Verilog hardware description language. Four internal 1-bit state registers are utilized to implement total 10 states. Alongside the positive triggered asynchronous reset and clock signals, control unit also receives 3-bit instruction bus and 1-bit output signal from the overflow flag as its inputs. Figure 6 represents the finite state diagram of the proposed control unit.

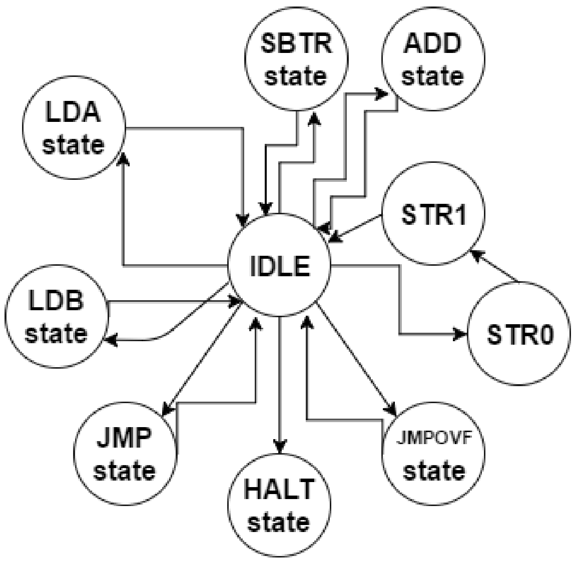


Figure 6 Finite state diagram of the control unit. Triggering control signals are omitted.

In order to directly perform block code mapping to obtain a fault tolerant circuit, one would require 22 bits to encode 17 bits of logic. However, this will cause decoder circuit to be large enough to decode 4 megabytes of decoding information, which will be infeasible to realize. To overcome this problem, 8-bit data buses x, y, and s are transformed to the 12-bit block code domain separately and we implemented our mapping algorithm to three data buses. The resulting circuit has 37 lines and 100 gates with 288 quantum cost while the original circuits had 32 gates with 64 cost. In the implementation, the gates with identical control lines are merged to reduce actual CMOS cost. The design is also verified in Xilinx ISE using a Verilog test fixture code.

The processor is synthesized in Cadence Encounter RTL Compiler using ams 0.35µm C35B4 process digital core library. From the RTL schematic, it can be observed that the total power consumption of the microprocessor is $1.68\,mW$. Examining the timing report, the maximum fanout is 5, maximum load capacitance is $74.8\,fF$, the maximum slew is $2248\,ps$, and the maximum delay is $1039\,ps$. The total area is $0.046\,mm^2$.

4. Simulation and Verification

After the analysis of the synthesized we simulated the microprocessor using the test code in Table 3. A Verilog test fixture code which initializes and maintains the reset and clock signals is executed in ISim simulator which is embedded inside Xilinx ISE. Contents of the memory locations during the execution of the program whose waveform pattern is given in Figure 7 are provided as follows:

Table 3 Test code for the microprocessor.

Memory Location	Opcode	Memory Content
R[0]	LDA	0100 0000
R[1]	#10	0000 1010
R[2]	LDB	0110 0000
R[3]	#5	0000 0101
R[4]	ADD	0000 0000
R[5]	STR	1000 0000
R[6]	@8	0000 1000
R[7]	HALT	11100000
R[8]	***	0000 0000
R[8]+	#15	0000 1111

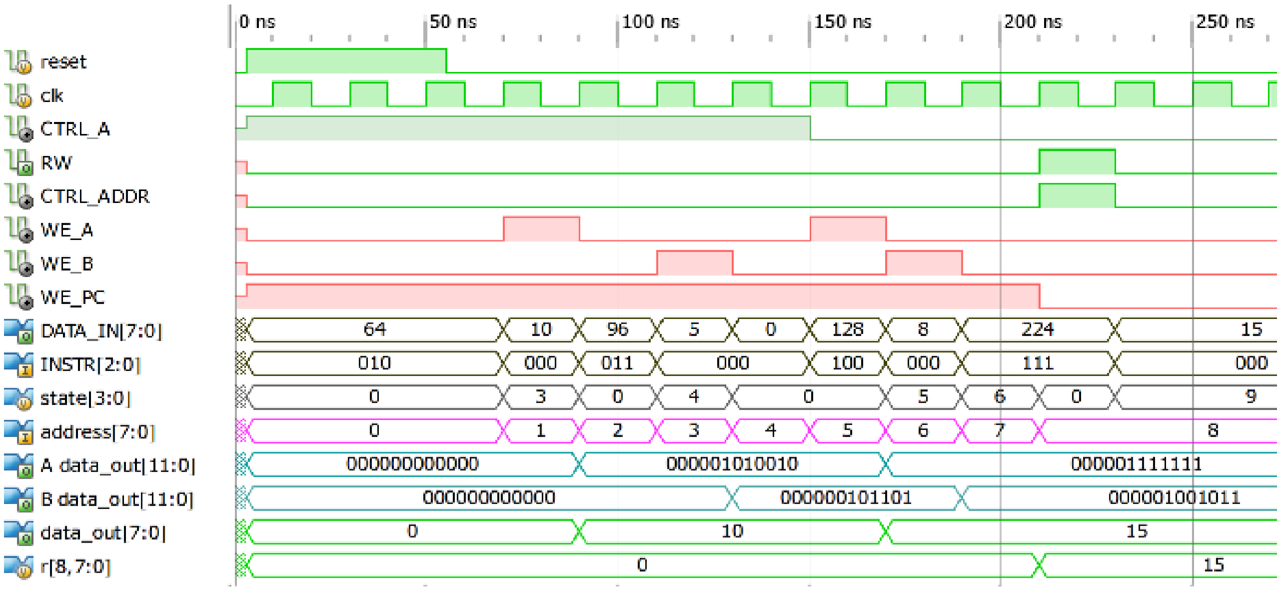


Figure 7 RTL simulation results for a test program.

After the program in Table 3 is executed, the resulting value 15 is written to the memory location R[13] while its updated value is denoted as $R[13]^{+}$. Memory contents can be followed at DATA_IN bus of Figure 7 in decimal radix. After the execution is done, the microprocessor performs a transition to the HALT state, thus no further opcode will be executed.

Upon the verification of the design, technology dependent post-synthesis Verilog code is modified to avoid logic trimming of our block code mapped fault tolerant reversible full adder. Alongside the design constraints (SDC) file, the resulting Verilog design is transferred to the Cadence Encounter Place and Route environment. The design was floorplanned and power routing by adding cell rings was performed. After adding three power stripes, end capacitors are added in order to achieve decoupling, i.e. electrically separating the substrate and wells by limiting the resistance in between them. After this step, cell blocks and IO pins are placed. No special handling for clock tree is performed since the expected clock delay is not critical considering the planned operating frequency. Afterwards, unused spaces are being filled with dummy metal. Finally, the design is automatically routed, generating 0 violations.

After the layout is obtained, post-layout report is examined. The resulting chip has 42 pins which consists of 12 DATA_IN, 12 DATA_OUT, 12 ADDRESS_OUT, 1 CLK, 1 RESET, 3 INSTR signals. Total number of utilized standard cells is 628, including left and right hand side end capacitors and fill shapes. The report also extracted the floating outputs (s and y) of ALU, which are generated as a sequence of reversibility. There are four routing layers MET1, MET2, MET3, and MET4.

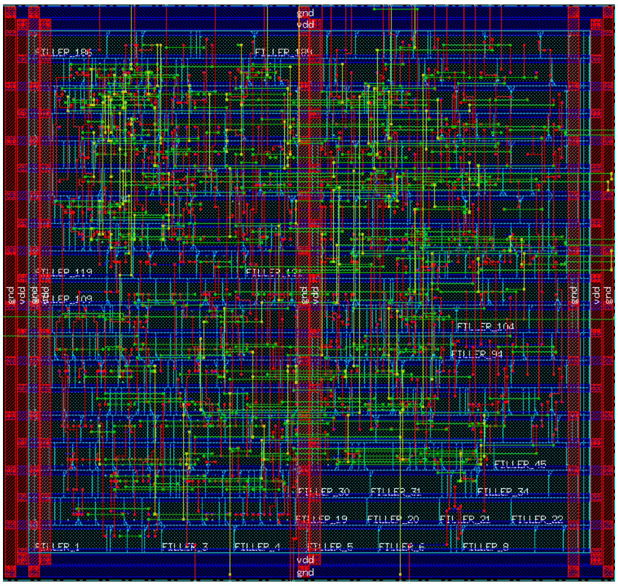


Figure 8 Encounter post-place-and-route view of the microprocessor.

The post-place-and-route and the layout pictures of the proposed microprocessor are given in Figure 8 and Figure 9, respectively. The whole chip occupies 0.08 mm^2 area. In the standard cell mapping, 19 rows are used with a gate density of 59.99% excluding physical cells. The core utilization of the whole chip is 83.22% in terms of standard cells, IO, and macro blocks. The total wire length is 23981.625 μm with an average wire length of 43.2881 μm per net. The detailed area numbers are given in Table 4.

Table 4. Detailed area report of the microprocessor.

Total area of Standard cells:	67085.200 μm^2
Total area of Standard cells (Subtracting Physical Cells):	40276.600 μm^2
Total area of Core:	67128.425 μm^2
Total area of Chip:	**80614.000 μm^2**

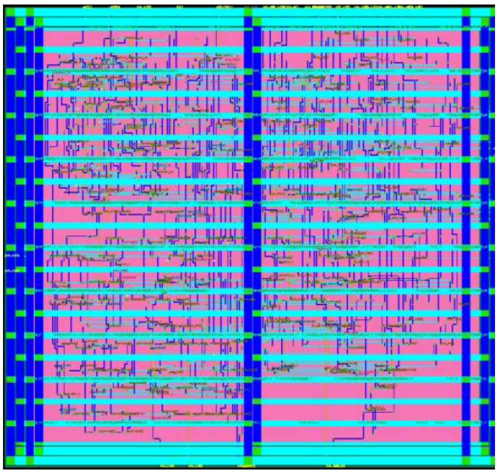


Figure 9 Negative GDSII layout of the microprocessor.

5. Conclusions

In this study, a fault-tolerant and defect-aware reversible RISC microprocessor has been designed as the proof of the concept where we show that reversible computing can be utilized to achieve perfect defect awareness. The design will be sent to fabrication to conduct further tests on the dual in-line packed tapeout. Prior to the submission of the design for tapeout, the ALU will be verified for a complete set of possible inputs. Multiplexers and control logic will be made reversible and/or defect-aware. The design will be carried out further design-rule checks. Post-layout simulation will be conducted including the delays resulting from parasitic capacitances and resistances.

6. Acknowledgments

This work is supported by The Scientific and Technological Research Council of Turkey (TUBITAK) 1002 Project # 215E268.

7. References

[1] Binder D., Smith, E.C. and Holman, A.B., 1975. “Satellite anomalies from galactic cosmic rays”, IEEE Trans. Nucl. Sci., vol. NS-22, no. 6, pp. 2675–2680.

[2] May, T.C. and Woods, M.H., 1978. “A new physical mechanism for soft errors in dynamic memories”, in Proc. Int’l Rel. Phys. Symp. (IRPS), pp. 33–40.

[3] Nicolaidis, M. ed., 2010. Soft errors in modern electronic systems (Vol. 41). Springer Science & Business Media.

[4] Toffoli, T., 1980. J. W. de Bakker and J. van Leeuwen, ed. "Reversible computing". Automata, Languages and Programming, Seventh Colloquium. Springer Verlag, Noordwijkerhout, Netherland.

[5] Maslov, D., Dueck, G. W., and Miller, D.M., “Toffoli Network Synthesis with Templates. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Vol. 24, No. 6, June 2005

[6] Brown, J., 2000, The Quest for the Quantum Computer, Touchstone, New York.

[7] Benioff, P., 1980. The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines. Journal of Statistical Physics, 22(5), pp.563-591

[8] Beckman, D., Chari, A.N., Devabhaktuni, S., and Preskill, J., 1996. Efficient networks for quantum factoring. Physical Review A, 54(2), p.1034.

[9] Moore, G.E., 1965. Cramming more components onto integrated circuits, Electronics Magazine. p. 4.

[10] Miller, D.M., Maslov, D., and Dueck, G.W., 2003, June. A transformation based algorithm for reversible logic synthesis. In Proceedings of the 40th annual Design Automation Conference (pp. 318-323). ACM

[11] Wille, R. and Drechsler, R., 2009, July. BDD-based synthesis of reversible logic for large functions. In Proceedings of the 46th Annual Design Automation Conference (pp. 270-275). ACM.

[12] Miller, D.M. and Thornton, M.A., 2006, May. QMDD: A decision diagram structure for reversible and quantum circuits. In Multiple-Valued Logic, 2006. ISMVL 2006. 36th International Symposium on (pp. 30-30). IEEE.

[13] Soeken, M., Wille, R., Hilken, C., Przigoda, N., and Drechsler, R., 2012, January. Synthesis of reversible circuits with minimal lines for large functions. In Design Automation Conference (ASP-DAC), 2012 17th Asia and South Pacific (pp. 85-92). IEEE.

[14] Bennett, C.H., 2003. Notes on Landauer's principle, reversible computation, and Maxwell's Demon. Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics, 34(3), pp.501-510.

[15] Landauer, R., 1961. Irreversibility and heat generation in the computing process. IBM journal of research and development, 5(3), pp.183-191.

[16] Willingham, D.J. and Kale, I., 2008. Using positive feedback adiabatic logic to implement reversible Toffoli gates.

[17] Thomsen, M.K., 2012. Design of reversible logic circuits using standard cells: Standard cells and functional programming. Department of Computer Science, University of Copenhagen.

[18] Burignat, S., Thomsen, M.K., Klimczak, M., Olczak, M., and De Vos, A., 2011. Interfacing reversible pass-transistor CMOS chips with conventional restoring CMOS circuits. InReversible Computation (pp. 112-122). Springer Berlin Heidelberg.

[19] Susam, O. and Altun, M., 2016. Fast Synthesis of Reversible Circuits using a Sorting Algorithm and Optimization. Journal of Multiple-Valued Logic and Soft Computing, accepted for publication.

[20] Wootters, W., Zurek, W., 1982. "A Single Quantum Cannot be Cloned". Nature 299: 802–803.

[21] Wille, R., Große, D., Teuber, L., Dueck, G.W., and Drechsler, R., 2008, May. RevLib: An online resource for reversible functions and reversible circuits. In Multiple Valued Logic, 2008. ISMVL 2008. 38th International Symposium on (pp. 220-225). IEEE.

[22] Parhami, B., 2006, October. Fault-tolerant reversible circuits. In Signals, Systems and Computers, 2006. ACSSC'06. Fortieth Asilomar Conference on (pp. 1726-1729). IEEE.

[23] Boykin, P.O. and Roychowdhury, V.P., 2005, June. Reversible fault-tolerant logic. In Dependable Systems and Networks, 2005. DSN 2005. Proceedings. International Conference on (pp. 444-453). IEEE.

[24] Przigoda, N., Dueck, G., Wille, R. and Drechsler, R., 2016. Fault Detection in Parity Preserving Reversible Circuits.

[25] DiVincenzo, D.P. and Shor, P.W., 1996. Fault-tolerant error correction with efficient quantum codes. Physical review letters, 77(15), p.3260.

[26] Van Rentergem, Y. and De Vos, A., 2005. Optimal design of a reversible full adder. International Journal of Unconventional Computing, 1(4), p.339.

Note:

Mustafa Altun received his BSc and MSc degrees in electronics engineering at Istanbul Technical University in 2004 and 2007, respectively. He received his PhD degree in electrical engineering with a PhD minor in mathematics at the University of Minnesota in 2012. Since 2013, he has served as an assistant professor of electrical engineering at Istanbul Technical University. Dr. Altun runs the Emerging Circuits and Computation (ECC) Group in the same university. Dr. Altun has been served as a principal investigator/researcher of various projects including EU H2020 RISE, National Science Foundation of USA (NSF), TUBITAK Career, and TUBITAK University-Industry Collaboration projects. He is an author of more than 30 peer reviewed papers and a book chapter, and the recipient of the TUBITAK Success, TUBITAK Career, and Werner von Siemens Excellence awards.

M. Hüsrev Cılasun is currently an engineer at Aselsan A.Ş. His research interests include reversible and quantum circuits, FPGA/ASIC design, digital signal and image processing, and machine learning. He is author of several conference and journal papers on EEG processing and autonomous direction estimation, number theory, and robotics.

A POWER RESOLUTION FOR COST EFFECTIVE COMPENSATION AND HARMONIC SOURCE DETECTION IN SMART POWER GRIDS

Murat Erhan BALCI[1], Mehmet Hakan HOCAOGLU [2]

[1]Dept. of Electrical and Electronics Engineering, Balikesir University, Balıkesir, Turkey
[2] Dept. of Electronics Engineering, Gebze Technical University, Kocaeli, Turkey
mbalci@balikesir.edu.tr, hocaoglu@gtu.edu.tr

Abstract: *It is seen from the preliminary work* [1] *of this paper that in the literature there is a need for a power resolution, which can be utilized for (i) the direct provision of the optimum compensation capacitor's power and (ii) obtaining meaningful information on the detection of the harmonic producing loads. This paper proposed a power resolution that can be used to achieve both goals under nosinusoidal and unbalanced conditions in the smart power grids. First goal is important for cost effective unity power factor compensation including a basic capacitor and an active compensator. The second goal is required for the practical detection of harmonic producing loads by using the demand meters, which are employed to measure the powers for the energy billing of the consumers. The proposed power resolution is based on the separation of load current into orthogonal components as active, reactive, scattered conductance, scattered susceptance, unbalanced conductance and unbalanced susceptance currents, which are all related to the conductance and susceptance parameters of the load. To show that the proposed resolution attains its goals, the simulation and experimental based analysis are presented in this paper.*
Keywords: *Power resolutions, nonsinusoidal and unbalanced systems, compensation, harmonic source detection.*

1. Introduction

By the proliferation of a.c. in the transmission and distribution systems, apparent power was defined as the product of voltage and current rms values to size the system equipment and to be a measure for the system efficiency [1, 2]. Historically, the current of the system was divided into two parts: These are; active current, which transports the net energy from source to the load, and reactive current that is the remaining current component when the active part is subtracted from the total current. According to this resolution, apparent power was expressed as the vector sum of active and reactive powers, which flows due to the active and reactive currents, respectively [3]. The ratio of active and apparent powers is named as the power factor, and be utilized to measure the efficiency of the power systems. In addition, conventionally, classical single-phase apparent power is directly extended to three-phase systems by treating each phase individually [1, 4]. Thus, arithmetic apparent power, which is calculated as the arithmetic sum of each phase's apparent power, and vector apparent power, which is calculated as the vector sum of total active and total reactive powers of the system, were constituted for the three-phase systems.

Nevertheless, due to the fact that the classical apparent power and its resolution are defined under sinusoidal and balanced conditions, they did not attain their goals in the case of nonsinusoidal and/or unbalanced conditions. Consequently, a number of apparent power definitions and their resolutions have been proposed for nonsinusoidal single-phase [3, 5-7] and nonsinusoidal & unbalanced three-phase systems [1, 4] to fulfil the gap left out in the classical apparent power concept. However, it is seen from the preliminary work [1] of this paper that in the literature there is a need for a power resolution, which can be utilized for (i) the direct provision of the optimum compensation capacitor's power and (ii) obtaining meaningful information on the detection of the harmonic producing loads. First goal is important for cost effective unity power factor compensation including a basic capacitor and an active compensator. The second goal is useful for the practical detection of harmonic producing loads by using the demand meters, which are employed to measure the powers for the energy billing of the consumers.

As a result, in this paper, for unbalanced nonsinusoidal three-phase and three-line systems, a power resolution is proposed by considering the above mentioned two goals. It should be reminded that the single-phase case of the proposed resolution is interpreted and analysed for compensation in [6, 7]. In addition, this study is partly presented in [8, 9].

Received on: 16.02.2016
Accepted on: 16.05.2016

2. Proposed Power Resolution

In this section, the power resolution proposed in [6, 7] is extended from nonsinusoidal single- phase system to nonsinusoidal and unbalanced three-phase system without neutral line (three-phase and three-line system). Here, it should be underlined that line currents of the three-phase system are nonsinusoidal and unbalanced. However, nonsinusoidal voltages measured in the system has negligible unbalance due to the fact that the considered system has a strong utility side, which means a very low short circuit impedance at fundamental frequency. According to these considerations, in the first step, by expressing line (l=a,b,c) voltages and currents as;

$$\vec{v}(t)=\begin{bmatrix} v_a(t)\\ v_b(t)\\ v_c(t)\end{bmatrix}=\begin{bmatrix}\sum_n \sqrt{2}V_n sin(\omega_n t+\theta_{a,n})\\ \sum_n \sqrt{2}V_n sin(\omega_n t+\theta_{b,n})\\ \sum_n \sqrt{2}V_n sin(\omega_n t+\theta_{c,n})\end{bmatrix},$$
$$\vec{i}(t)=\begin{bmatrix} i_a(t)\\ i_b(t)\\ i_c(t)\end{bmatrix}=\begin{bmatrix}\sum_n \sqrt{2}I_{a,n} sin(\omega_n t+\delta_{a,n})\\ \sum_n \sqrt{2}I_{b,n} sin(\omega_n t+\delta_{b,n})\\ \sum_n \sqrt{2}I_{c,n} sin(\omega_n t+\delta_{c,n})\end{bmatrix} \quad (1)$$

the balanced and unbalanced parts of the line currents can be separated as below:

$$\vec{i}(t)=\vec{i}_B(t)+\vec{i}_u(t) \quad (2)$$

To find the expression of balanced current component ($\vec{i}_B(t)$), the powers ($U_{l,n}$), which are drawn due to the n*th* harmonic line currents in phase with the n*th* harmonic of respective line voltages, and the powers ($Q_{l,n}$), which are drawn due to the n*th* harmonic line currents in quadrature with the n*th* harmonic of respective line voltages, are calculated:

$$U_{l,n}=V_n I_{l,n} cos(\theta_{l,n}-\delta_{l,n}) \qquad (l=a,b,c) \quad (3)$$
$$Q_{l,n}=V_n I_{l,n} sin(\theta_{l,n}-\delta_{l,n}) \qquad (l=a,b,c) \quad (4)$$

And then, these powers are shared to each phase equally; thus, fictitious n*th* harmonic balanced active (P_n) and reactive (Q_n) powers are found to be:

$$P_n=\frac{1}{3}(U_{a,n}+U_{b,n}+U_{c,n}) \quad (5)$$

$$Q_n=\frac{1}{3}(Q_{a,n}+Q_{b,n}+Q_{c,n}) \quad (6)$$

For n*th* harmonic, balanced active and balanced reactive powers, given in (5) and (6), are drawn by the balanced part of n*th* harmonic equivalent impedance of the three-phase load. Thus, for each phase of the load, n*th* harmonic balanced conductance ($G_{B,n}$) and n*th* harmonic balanced susceptance ($B_{B,n}$) can be calculated as;

$$G_{B,n}=\frac{P_n}{V_n^2} \quad (7)$$
$$B_{B,n}=\frac{Q_n}{V_n^2} \quad (8)$$

By using n*th* harmonic balanced conductance and n*th* harmonic balanced susceptance parameters, the balanced current component can be expressed as;

$$\vec{i}_B(t)=\begin{bmatrix}\sum_n \sqrt{2}G_{B,n}V_n\, sin(\omega_n t+\theta_{a,n})\\ \sum_n \sqrt{2}G_{B,n}V_n\, sin(\omega_n t+\theta_{b,n})\\ \sum_n \sqrt{2}G_{B,n}V_n\, sin(\omega_n t+\theta_{c,n})\end{bmatrix}+\begin{bmatrix}\sum_n \sqrt{2}B_{B,n}V_n\, sin\left(\omega_n t+\theta_{a,n}-\frac{\pi}{2}\right)\\ \sum_n \sqrt{2}B_{B,n}V_n\, sin\left(\omega_n t+\theta_{b,n}-\frac{\pi}{2}\right)\\ \sum_n \sqrt{2}B_{B,n}V_n\, sin\left(\omega_n t+\theta_{c,n}-\frac{\pi}{2}\right)\end{bmatrix} \quad (9)$$

To find an expression for unbalanced current component ($\vec{i}_u(t)$), the n*th* harmonic conductance ($G_{l,n}$) and n*th* harmonic susceptance ($B_{l,n}$) values seen from *l*=a, b and c phases of the load are written as:

$$G_{l,n}=\frac{U_{l,n}}{V_n^2} \quad (10)$$
$$B_{l,n}=\frac{Q_{l,n}}{V_n^2} \quad (11)$$

Thus, for each phase of the load, n*th* harmonic unbalanced conductance ($G_{l,n}^u$) and n*th* harmonic unbalanced susceptance ($B_{l,n}^u$) values can be found as;

$$G_{l,n}^u=G_{l,n}-G_{B,n} \quad (12)$$
$$B_{l,n}^u=B_{l,n}-B_{B,n} \quad (13)$$

By using $G_{l,n}^u$ and $B_{l,n}^u$, the unbalanced current ($\vec{i}_u(t)$) can be divided into two parts as unbalanced in-phase ($\vec{i}_{up}(t)$) and unbalanced quadrature ($\vec{i}_{uq}(t)$) currents, respectively:

$$\vec{i}_u(t)=\vec{i}_{up}(t)+\vec{i}_{uq}(t)$$
$$=\begin{bmatrix}\sum_n \sqrt{2}G^u_{a,n}V_n\, sin\left(\omega_n t+\theta_{a,n}\right)\\ \sum_n \sqrt{2}G^u_{b,n}V_n\, sin\left(\omega_n t+\theta_{b,n}\right)\\ \sum_n \sqrt{2}G^u_{c,n}V_n\, sin\left(\omega_n t+\theta_{c,n}\right)\end{bmatrix}+$$
$$+\begin{bmatrix}\sum_n \sqrt{2}B^u_{a,n}V_n\, sin\left(\omega_n t+\theta_{a,n}-\frac{\pi}{2}\right)\\ \sum_n \sqrt{2}B^u_{b,n}V_n\, sin\left(\omega_n t+\theta_{b,n}-\frac{\pi}{2}\right)\\ \sum_n \sqrt{2}B^u_{c,n}V_n\, sin\left(\omega_n t+\theta_{c,n}-\frac{\pi}{2}\right)\end{bmatrix} \quad (14)$$

In the second step, the balanced current ($\vec{i}_B(t)$) is decomposed into four orthogonal components namely; active, reactive, scattered conductance and scattered susceptance currents by treating each phase individually. The methodology presented in [6, 7] is considered for the decomposition. This is valid due to the fact that voltage has negligible unbalance and the decomposed current part is completely balanced.

Accordingly, for each phase of the load, equivalent conductance (G_e) is calculated as;

$$G_e=\frac{\sum_n P_n}{\sum_n V_n^2} \quad (15)$$

and active current is written in terms of G_e:

$$\vec{i}_{ac}(t)=\begin{bmatrix}\sum_n \sqrt{2}G_eV_n\, sin\left(\omega_n t+\theta_{a,n}\right)\\ \sum_n \sqrt{2}G_eV_n\, sin\left(\omega_n t+\theta_{b,n}\right)\\ \sum_n \sqrt{2}G_eV_n\, sin\left(\omega_n t+\theta_{c,n}\right)\end{bmatrix} \quad (16)$$

And then, for each phase of the load, by calculating n*th* harmonic equivalent susceptance ($B_{e,n}$) as;

$$B_{e,n}=nB_{e,1}=n\frac{1}{Xc_1},\; Xc_1=\sum_n n^2V_n^2 \Big/ \sum_n nQ_n \quad (17)$$

reactive current can be expressed as:

$$\vec{i}_r(t)=\begin{bmatrix}\sum_n \sqrt{2}B_{e,n}V_n\, sin\left(\omega_n t+\theta_{a,n}-\frac{\pi}{2}\right)\\ \sum_n \sqrt{2}B_{e,n}V_n\, sin\left(\omega_n t+\theta_{b,n}-\frac{\pi}{2}\right)\\ \sum_n \sqrt{2}B_{e,n}V_n\, sin\left(\omega_n t+\theta_{c,n}-\frac{\pi}{2}\right)\end{bmatrix} \quad (18)$$

Sequently, using (16) and (18), the expression of the balanced current is rewritten in terms of active, reactive and scattered ($\vec{i}_s(t)$) currents:

$$\vec{i}_B(t)=\begin{bmatrix}\sum_n \sqrt{2}G_eV_n\, sin\left(\omega_n t+\theta_{a,n}\right)\\ \sum_n \sqrt{2}G_eV_n\, sin\left(\omega_n t+\theta_{b,n}\right)\\ \sum_n \sqrt{2}G_eV_n\, sin\left(\omega_n t+\theta_{c,n}\right)\end{bmatrix}+$$
$$+\begin{bmatrix}\sum_n \sqrt{2}B_{e,n}V_n\, sin\left(\omega_n t+\theta_{a,n}-\frac{\pi}{2}\right)\\ \sum_n \sqrt{2}B_{e,n}V_n\, sin\left(\omega_n t+\theta_{b,n}-\frac{\pi}{2}\right)\\ \sum_n \sqrt{2}B_{e,n}V_n\, sin\left(\omega_n t+\theta_{c,n}-\frac{\pi}{2}\right)\end{bmatrix}+\vec{i}_s(t) \quad (19)$$

By equating the right hand sides of (9) and (19), $\vec{i}_s(t)$ can be expressed as:

$$\vec{i}_s(t)=\begin{bmatrix}\sum_n \sqrt{2}\left(G_{B,n}-G_e\right)V_n\, sin\left(\omega_n t+\theta_{a,n}\right)\\ \sum_n \sqrt{2}\left(G_{B,n}-G_e\right)V_n\, sin\left(\omega_n t+\theta_{b,n}\right)\\ \sum_n \sqrt{2}\left(G_{B,n}-G_e\right)V_n\, sin\left(\omega_n t+\theta_{c,n}\right)\end{bmatrix}+$$
$$+\begin{bmatrix}\sum_n \sqrt{2}\left(B_{B,n}-B_{e,n}\right)V_n\, sin\left(\omega_n t+\theta_{a,n}-\frac{\pi}{2}\right)\\ \sum_n \sqrt{2}\left(B_{B,n}-B_{e,n}\right)V_n\, sin\left(\omega_n t+\theta_{b,n}-\frac{\pi}{2}\right)\\ \sum_n \sqrt{2}\left(B_{B,n}-B_{e,n}\right)V_n\, sin\left(\omega_n t+\theta_{c,n}-\frac{\pi}{2}\right)\end{bmatrix} \quad (20)$$

And lastly, the parts related to conductance and susceptance of $\vec{i}_s(t)$ can be named as scattered conductance current;

$$\vec{i}_{sc}(t)=\begin{bmatrix}\sum_n \sqrt{2}\left(G_{B,n}-G_e\right)V_n\, sin\left(\omega_n t+\theta_{a,n}\right)\\ \sum_n \sqrt{2}\left(G_{B,n}-G_e\right)V_n\, sin\left(\omega_n t+\theta_{b,n}\right)\\ \sum_n \sqrt{2}\left(G_{B,n}-G_e\right)V_n\, sin\left(\omega_n t+\theta_{c,n}\right)\end{bmatrix} \quad (21)$$

and scattered susceptance current;

$$\vec{i}_{ss}(t)=\begin{bmatrix}\sum_n \sqrt{2}\left(B_{B,n}-B_{e,n}\right)V_n\, sin\left(\omega_n t+\theta_{a,n}-\frac{\pi}{2}\right)\\ \sum_n \sqrt{2}\left(B_{B,n}-B_{e,n}\right)V_n\, sin\left(\omega_n t+\theta_{b,n}-\frac{\pi}{2}\right)\\ \sum_n \sqrt{2}\left(B_{B,n}-B_{e,n}\right)V_n\, sin\left(\omega_n t+\theta_{c,n}-\frac{\pi}{2}\right)\end{bmatrix} \quad (22)$$

As a result, the load current can be written as;

$$\vec{i}\left(t\right)=\vec{i}_{ac}\left(t\right)+\vec{i}_{r}\left(t\right)+\vec{i}_{sc}\left(t\right)+\vec{i}_{ss}\left(t\right)+\vec{i}_{up}\left(t\right)+\vec{i}_{uq}\left(t\right) \quad (23)$$

By considering the collective rms current expression of the Buchollz's apparent power definition [1, 4, 10], the rms values of the currents placed in (23) can be calculated as;
total current's rms value;

$$I_{\Sigma}=\sqrt{\frac{1}{T}\int_{0}^{T}\left(\vec{i}\left(t\right)\right)^{Tr}\cdot\vec{i}\left(t\right)dt}=\sqrt{\sum_{l} I_{l}^{2}} \quad (24)$$

active current's rms value,

$$I_{ac}=\sqrt{\frac{1}{T}\int_{0}^{T}\left(\vec{i}_{ac}\left(t\right)\right)^{Tr}\cdot\vec{i}_{ac}\left(t\right)dt}=\sqrt{3G_{e}^{2}\sum_{n} V_{n}^{2}} \quad (25)$$

reactive current's rms value,

$$I_{r}=\sqrt{\frac{1}{T}\int_{0}^{T}\left(\vec{i}_{r}\left(t\right)\right)^{Tr}\cdot\vec{i}_{r}\left(t\right)dt}=\sqrt{3\sum_{n} B_{e,n}^{2}V_{n}^{2}} \quad (26)$$

scattered conductance current's rms value,

$$\begin{aligned}I_{sc}&=\sqrt{\frac{1}{T}\int_{0}^{T}\left(\vec{i}_{sc}\left(t\right)\right)^{Tr}\cdot\vec{i}_{sc}\left(t\right)dt}\\&=\sqrt{3\sum_{n}\left(G_{B,n}-G_{e}\right)^{2}V_{n}^{2}}\end{aligned} \quad (27)$$

scattered susceptance current's rms value,

$$\begin{aligned}I_{ss}&=\sqrt{\frac{1}{T}\int_{0}^{T}\left(\vec{i}_{ss}\left(t\right)\right)^{Tr}\cdot\vec{i}_{ss}\left(t\right)dt}\\&=\sqrt{3\sum_{n}\left(B_{B,n}-B_{e,n}\right)^{2}V_{n}^{2}}\end{aligned} \quad (28)$$

unbalanced in phase current's rms value,

$$\begin{aligned}I_{up}&=\sqrt{\frac{1}{T}\int_{0}^{T}\left(\vec{i}_{up}\left(t\right)\right)^{Tr}\cdot\vec{i}_{up}\left(t\right)dt}\\&=\sqrt{\sum_{n}\left[\left(G_{a,n}^{u}\right)^{2}+\left(G_{b,n}^{u}\right)^{2}+\left(G_{c,n}^{u}\right)^{2}\right]V_{n}^{2}}\end{aligned} \quad (29)$$

and unbalanced quadrature current's rms value,

$$\begin{aligned}I_{uq}&=\sqrt{\frac{1}{T}\int_{0}^{T}\left(\vec{i}_{uq}\left(t\right)\right)^{Tr}\cdot\vec{i}_{uq}\left(t\right)dt}\\&=\sqrt{\sum_{n}\left[\left(B_{a,n}^{u}\right)^{2}+\left(B_{b,n}^{u}\right)^{2}+\left(B_{c,n}^{u}\right)^{2}\right]V_{n}^{2}}\end{aligned} \quad (30)$$

In (24)-(30), the superscript "Tr" denotes transpose of the respected three-phase current vector.
Since n*th* harmonic of conductance based current components are in phase with the n*th* harmonic of voltage and the n*th* harmonic of susceptance based current components are in quadrature with the n*th* harmonic of voltage, all combinations between conductance based currents and susceptance based currents are orthogonal. The orthogonality proofs of the rest combinations of the proposed current components are provided in [8]. Therefore, the collective rms value of the total current can be expressed as the vector sum of the collective rms values of the proposed current components:

$$I_{\Sigma}^{2}=I_{ac}^{2}+I_{r}^{2}+I_{sc}^{2}+I_{ss}^{2}+I_{up}^{2}+I_{uq}^{2} \quad (31)$$

Finally, if both sides of (31) are multiplied by the square of collective voltage rms value ($V_{\Sigma}^{2}=\sum_{l} V_{l}^{2}$), the resolution of Buchollz's apparent power (S) can be obtained as;

$$\begin{aligned}S^{2}&=V_{\Sigma}^{2}I_{\Sigma}^{2}=V_{\Sigma}^{2}I_{ac}^{2}+V_{\Sigma}^{2}I_{r}^{2}+V_{\Sigma}^{2}I_{sc}^{2}+V_{\Sigma}^{2}I_{ss}^{2}+V_{\Sigma}^{2}I_{up}^{2}+V_{\Sigma}^{2}I_{uq}^{2}\\&=P^{2}+Q_{r}^{2}+D_{sc}^{2}+D_{ss}^{2}+D_{up}^{2}+D_{uq}^{2}\end{aligned} \quad (32)$$

In (32), power components are named as active (P), reactive (Q_r), scattered conductance (D_{sc}), scattered susceptance (D_{ss}), unbalanced in phase (D_{up}) and unbalanced quadrature (D_{uq}) powers. Note that for the systems without zero sequence voltages, the apparent powers of Bucholz [1, 10] and IEEE standard 1459 [11] have the same numerical values [12, 13]. Thus, proposed power resolution also decomposes IEEE standard 1459 apparent power.

To point out the novelty of the proposed resolution, here it is compared with Czarnecki's power resolution [14, 15], which is similar to the proposed one:

- The reactive power component of Czarnecki's resolution is divided into two power components, namely; reactive power and scattered susceptance power in the proposed resolution. The reactive power component of the proposed resolution is completely compensated when the power factor is maximized by the balanced three-phase capacitors bank. However, this is not the case for the reactive power component of Czarnecki's resolution.
- And also, the unbalanced power component of Czarnecki's power resolution is decomposed into unbalanced in phase and unbalanced quadrature powers in the proposed resolution:

$$D_{u}^{2}=D_{up}^{2}+D_{uq}^{2}=V_{\Sigma}^{2}\left(I_{up}^{2}+I_{uq}^{2}\right) \quad (33)$$

As a result, in the proposed resolution, all power components are expressed in terms of the conductance and susceptance parameters of the load. This should be useful for cooperative design and control of different types of harmonic, unbalance and reactive power compensators.

The contributions of the proposed power resolution to frequency-domain power theory are detailed below:

2.1. Providing a Tool for Cost Effective Unity Power Factor Compensation

As mentioned before, the reactive power (Q_r) of the proposed resolution is completely compensated when the power factor is maximized by the balanced three-phase capacitors bank. But, its other nonactive powers (D_{sc}, D_{ss}, D_{up} and D_{uq}) do not have any portions compensable via the balanced capacitors bank. This means that Q_r gives the power of optimum balanced capacitive compensator (S_{pC}), which maximizes the power factor under the nonsinusoidal voltage and current conditions:

$$S_{pC} = Q_r \tag{34}$$

In addition to that, for the cost effective unity power factor compensation scheme including the balanced capacitors bank and an active filter, the power (S_{aC}) of the active filter can practically be sized as the vector sum of D_{sc}, D_{ss}, D_{up} and D_{uq}:

$$S_{aC} = \sqrt{D_{sc}^2 + D_{ss}^2 + D_{up}^2 + D_{uq}^2} \tag{35}$$

For this compensation scheme, according to unity power factor instantaneous compensation strategy [16], the active filter (or compensator) should inject the current ($\vec{i}_{af}(t)$), given in (36), into the system:

$$\vec{i}_{af}(t) = \vec{i}_{sc}(t) + \vec{i}_{ss}(t) + \vec{i}_{up}(t) + \vec{i}_{uq}(t) \tag{36}$$

2.2. Providing a Tool for the Detection of the Harmonic Producing Loads

From the outlines of the proposed power resolution, it can be qualitatively concluded that two power components could be used to detect the harmonic producing loads: These are; scattered conductance power (D_{sc}), which occurs by the difference between nth harmonic balanced conductance and equivalent conductance, and scattered susceptance power (D_{ss}), which occurs by the difference between nth harmonic balanced susceptance and nth harmonic equivalent susceptance. However, D_{ss} is quiet sensitive to the source side's harmonic distortion, thus; its usage for the detection of the harmonic producing loads will be problematic. On the other hand, D_{sc} has the cases underlined below;

- **Sinusoidal Voltage (or Voltage with Negligible THD_V):** For a linear load under negligible voltage distortion, apparent power is very close to the fundamental harmonic apparent power due to the fact that both voltage and current harmonics have negligible magnitudes. As a result, for the same load-voltage case, D_{sc} has negligible value. On the contrary, for a nonlinear (or harmonic producing) load under the same voltage case, the load's balanced conductances calculated for the harmonics a part from fundamental one have the considerable values due to the fact that the load injects current harmonics, which is extremely higher than the respective voltage harmonics. Therefore, the rms value of the scattered conductance current (I_{sc}) and D_{sc} have considerably high values for the nonlinear load condition.
- **Nonsinusoidal Voltage:** In IEEE standard 519 [17], the maximum permissible value of THD_V is determined as 5% at the bus voltages lower than 69kV, 2.5% at the bus voltages between 69kV and 161kV and 1.5% at the bus voltages higher than 161kV. Thus, according to the same standard, the maximum value of THD_V measured at point of common coupling (pcc) can be assumed as 5%. For the maximum voltage distortion level, a linear load draws very small harmonic currents if there is no any resonance condition in the system. As a result, I_{sc} and D_{sc} of the linear load will be small. For a harmonic producing load under the same voltage distortion level, it is feasible that its current have some harmonic components, which is extremely larger than respective harmonic components of the voltage, in other words it behaves as considerably large conductances for the respective harmonic numbers. Consequently, I_{sc} and D_{sc} drawn by the harmonic producing load will have considerable values under the distorted supply voltage condition.

According to the manners mentioned above, it can be concluded whether the load has a non-harmonic producing characteristic or not.

3. Analysis Results on Reactive Power Compensation

In this section, it will be demonstrated that the proposed resolution can be used as a tool to obtain the power of optimum balanced capacitive compensator, which maximizes power factor under nonsinusoidal conditions. In addition, it will also be pointed out that the unity power factor can be achieved cost effectively by using the combination of the optimum balanced capacitive and active compensators (hybrid compensation). In other words, it will be shown that the power of active compensator in the hybrid compensation scheme could be smaller than used in pure active compensation scheme. The merit of the proposed resolution for the direct implementation of hybrid compensation will be simulated in the system, given in Figure 1. The simulated system consists of four kinds of loads, which are a resistive load supplied with frequency converter, a six-pulse rectifier with resistive load, a six-pulse rectifier with dc motor and an unbalanced inductive load. For the system without compensation (NC), the waveshapes of the line voltages and line currents are given in Figure 2 and Figure 3, respectively.

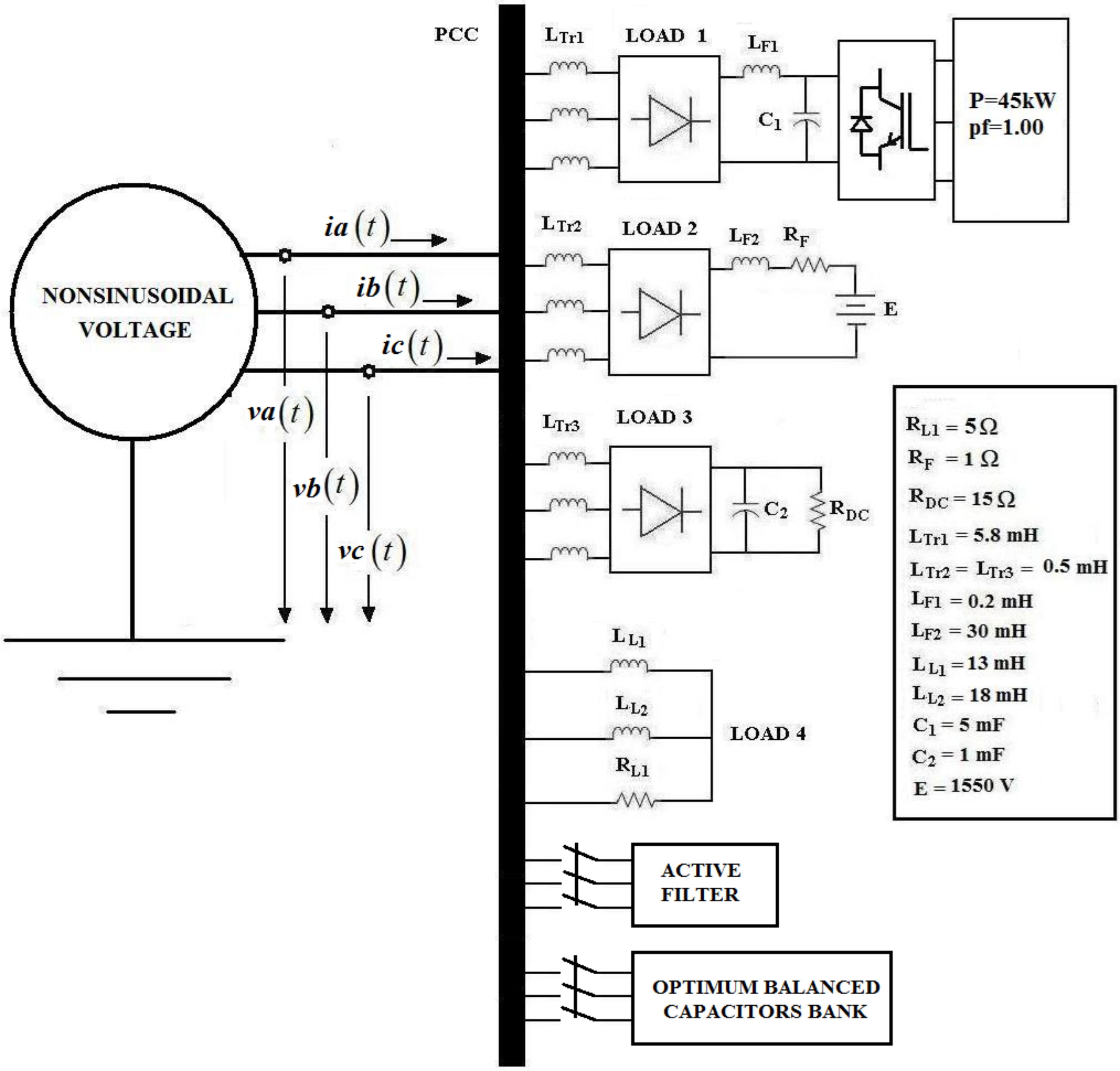


Figure 1. The simulated system considered for the compensation analysis

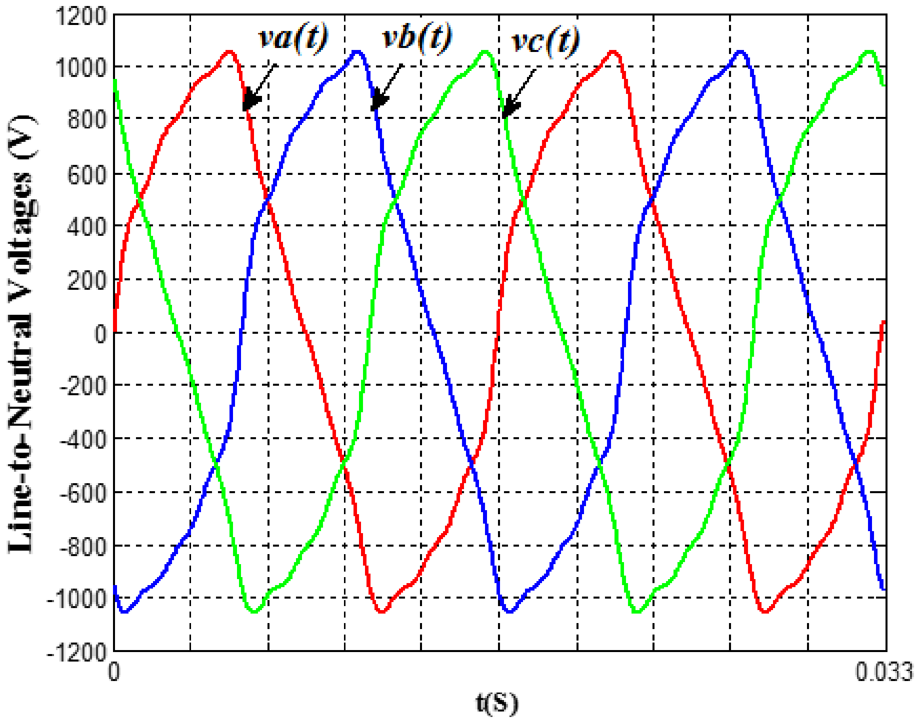


Figure 2. The wave shapes of the line voltages of the simulated system

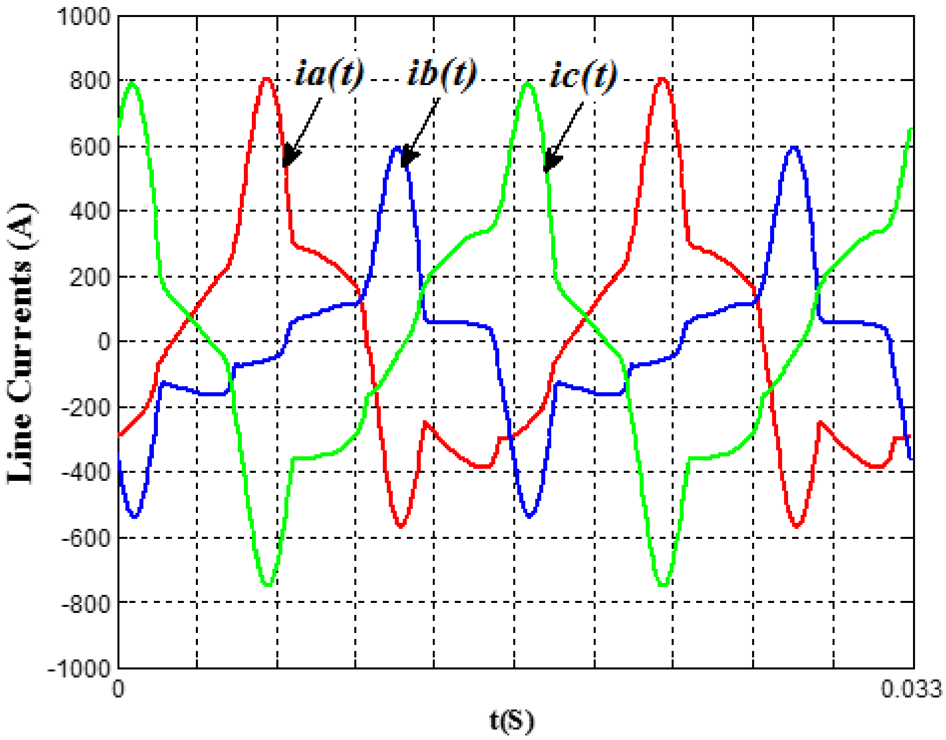


Figure 3. The wave shapes of the line currents of the simulated system without compensation (NC)

Figure 2 shows that line voltages have nonsinusoidal and balanced wave forms, which have 10% of THD_V. The line currents, plotted in Figure 3, have nonsinusoidal wave forms with 42%, 70% and 39% of THD_I, respectively. They are also significantly unbalanced: the ratio between fundamental harmonic negative and fundamental harmonic positive sequence magnitudes of the line currents, I_1^-/I_1^+, is 32.3%.

For the system without compensation (NC), with optimum balanced capacitive compensation (OBC), active compensation (AC) and hybrid compensation (HC), the proposed power components (P, Q_r, D_{sc}, D_{ss}, D_{up} and D_{uq}), apparent power (S), the powers of passive and active compensators (S_{pC} and S_{aC}) and power factor (pf=P/S) are plotted in Figure 4 and Figure 5. It should be noted that optimum balanced capacitive compensator is a star connected three identical capacitors, of which the capacitances can be determined using (17), active compensation is undertaken by unity power factor compensation strategy [16].

Figure 4 and 5 show that for NC case, pf, Q_r and S are 0.756, 0.435pu and 1.00pu. In addition to that, for NC case, vector sum of D_{sc}, D_{ss}, D_{up} and D_{uq} is 0.489pu. For OBC case, Q_r is completely compensated and pf is improved from 0.756 to 0.840 by using star connected three identical capacitors, of which the power (S_{pc}) is 0.435pu. Obviously, all power components except Q_r have the same values for NC and OBC cases. Thus, one can see that Q_r gives the power of optimum balanced capacitive compensator. On the other hand, pf is still smaller than unity for OBC since D_{sc}, D_{ss}, D_{up} and D_{uq} can not be compensated via the balanced capacitive compensator. For AC case, unity power factor is achieved by using only active compensator, of which power (S_{ac}) is equal to 0.654pu, and S is decreased to 0.756pu. For HC case, unity power factor is achieved by using the optimum balanced capacitive compensator, of which power is 0.435pu, and an active compensator, of which power is equal to 0.489pu. Therefore, it can be pointed out that the power of active compensator used in hybrid compensation is 74.7% of active compensator's power calculated for the pure active compensation. In addition, it should also be underlined that for HC case, Q_r and the vector sum of the D_{sc}, D_{ss}, D_{up} and D_{uq} are equal to S_{pC} and S_{aC}, respectively. This means that the proposed resolution can practically be employed as a tool to design the cost effective unity power factor compensator consisting of the basic capacitors and an active compensator.

4. Experimental Analysis on the Detection of Harmonic Producing Loads

In this section, the harmonic producing load detection method based on the proposed power resolution is statistically evaluated by using a real test system, which comprises various types of linear and nonlinear loads. The schematic of the system are depicted in Figure 6. In the schematic, PC processes voltage and current data, and controls the programmable power supply, which generates the desired voltage wave forms. A R-L impedance with X/R=0.5, an induction machine working with the constant speed and constant torque cases under full loading, a dimmer controlled R-L impedance (X/R=0.5 and the triac conduction angles: 90º-270º), a number of computers and a number of compact fluorescent lamps are the load types employed in the test system.

Each one of the loads are supplied with one sinusoidal and one hundred randomly produced different distorted voltages with 5% value of THD_V. For the sinusoidal excitation, voltage and current pairs of the loads are plotted in Figure 7. It is seen from Figure 7 (a) that the current of the R-L impedance load under sinusoidal supply voltage is sinusoidal. Figure 7 (b) and (c) shows that the induction machine draws a current with small amount of THD_I, which is measured as 5%, under sinusoidal supply voltage. On the other hand, the currents of the dimmer controlled R-L impedance, computers and compact fluorescent lamps, of which THD_I values are measured as 50, 185 and 115%, respectively, are seen as highly distorted from Figure 7 (d), (e) and (f).

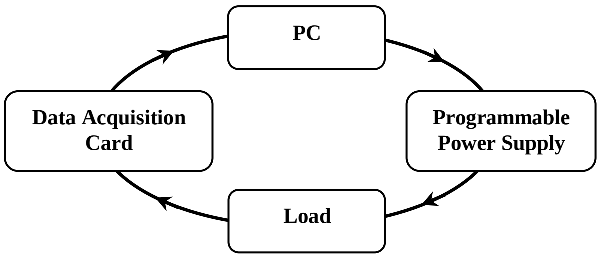


Figure 6. The schematic of the test system used for the harmonic source detection analysis

The normalised values of the powers measured under sinusoidal supply voltage and the histograms of the normalised values of the powers measured under distorted test voltages are presented in Figure 8 and Figure 9, respectively. Figure 8 (a) shows that the R-L impedance draws P and Q_r measured as 0.89 and 0.44 pu, respectively. In addition, D_{sc} and D_{ss} values of the R-L impedance are almost zero under the sinusoidal supply voltage. Figure 8 (a) and (d) point out that for the same load type the P, Q_r and D_{sc} values measured under one hundred distorted test voltages are very close to their measured values under sinusoidal supply voltage. However, under the distorted test voltages, the D_{ss} values drawn by the R-L impedance vary between 0.1 and 0.2 pu. This clearly means that D_{ss} is strongly dependent on the source side distortion.

Figure 8 (b) points out that the induction machine working with constant speed draws P, Q_r, D_{sc} and D_{ss} measured as 0.81, 0.58, 0.03 and 0.04 pu, respectively. It is seen from Figure 8 (b) and (e) that P, Q_r and D_{sc} values measured under the distorted test voltages are around their measured values under sinusoidal supply voltage. However, for the distorted test voltages, D_{ss} varies between 0.1 and 0.2 pu.

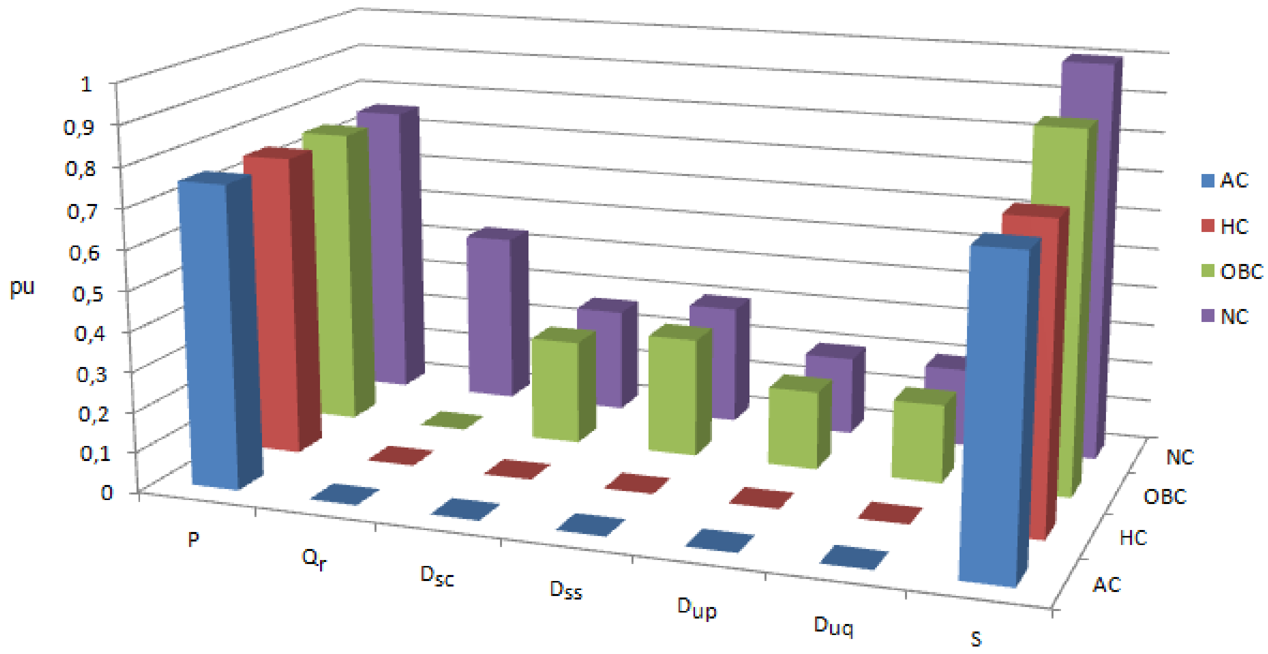


Figure 4. Powers measured at the pcc of the simulated system for the NC, OBC, HC and AC cases

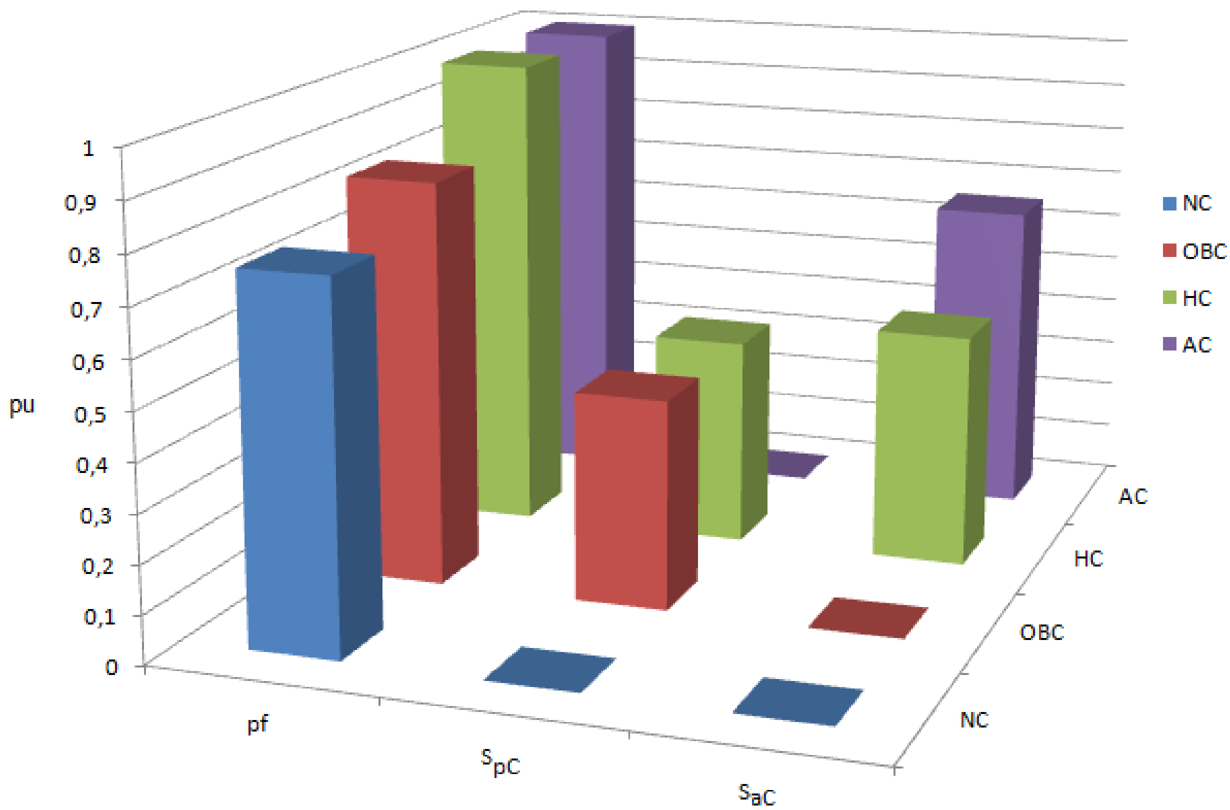


Figure 5. The powers of compensators and power factors for the NC, OBC, HC and AC cases

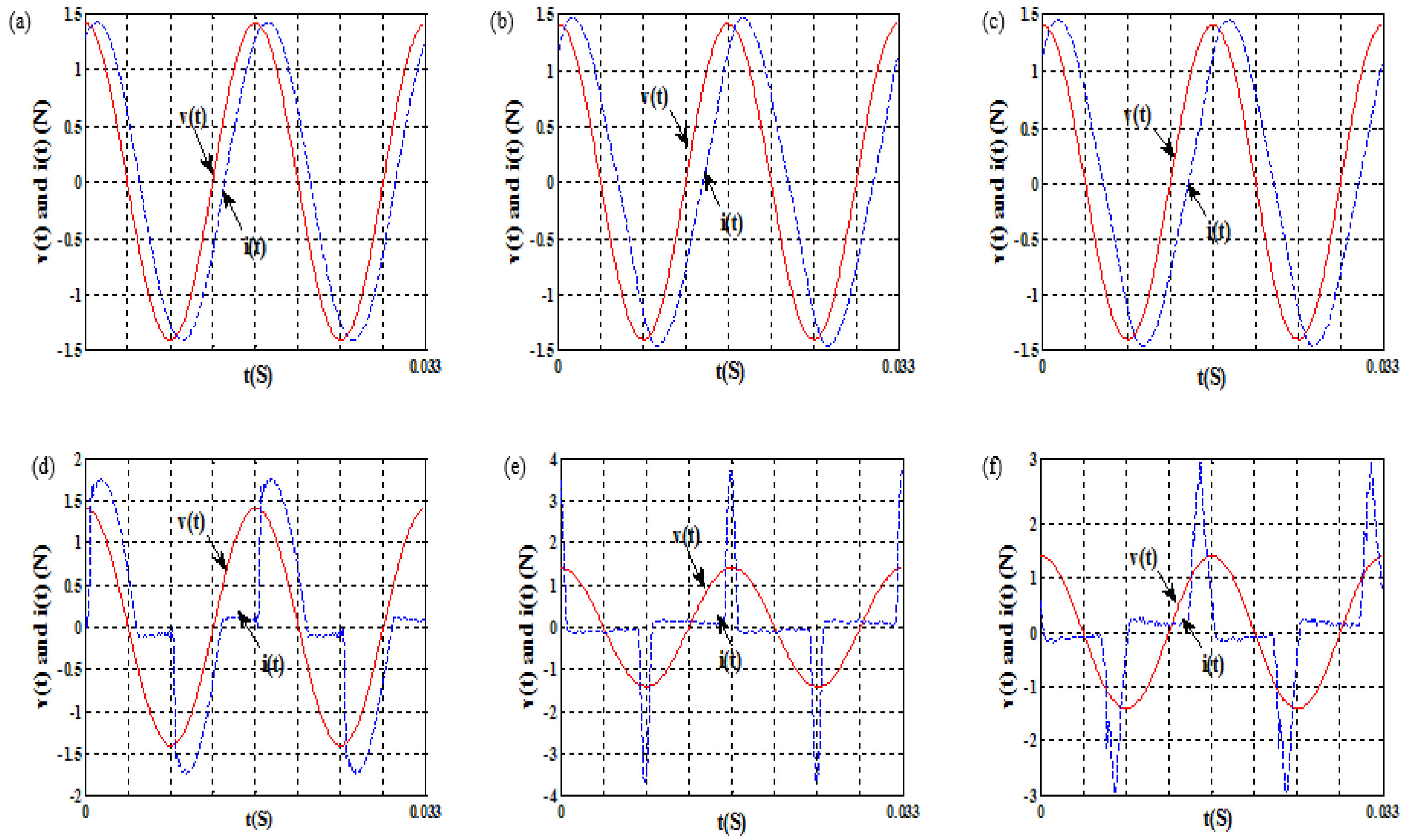


Figure 7. For sinusoidal excitation, the voltage and current pairs of (a) the R-L impedance, (b) the induction machine working with constant speed, (c) the induction machine working with constant torque, (d) the R-L impedance controlled with a dimming circuit, (e) a number of computers and (f) a number of compact fluorescent lamps.

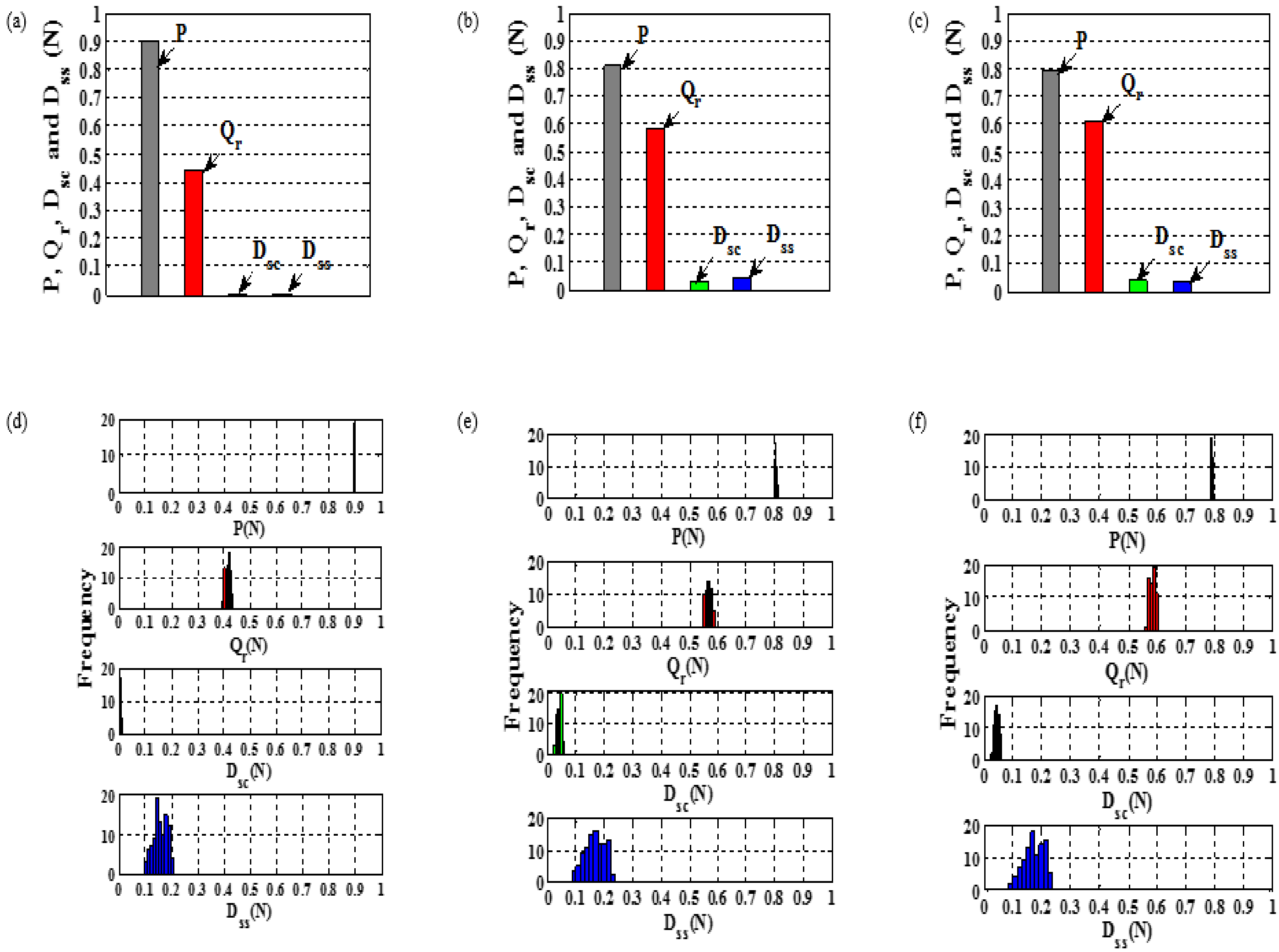


Figure 8. The normalised values of P, Q_r, D_{sc} and D_{ss} measured under sinusoidal supply voltage for (a) a R-L impedance, (b) the constant speed case of an induction machine and (c) the constant torque case of an induction machine, and the histograms of the normalised powers measured under distorted test voltages for (d) a R-L impedance, (e) the constant speed case of an induction machine and (f) constant torque case of an induction machine.

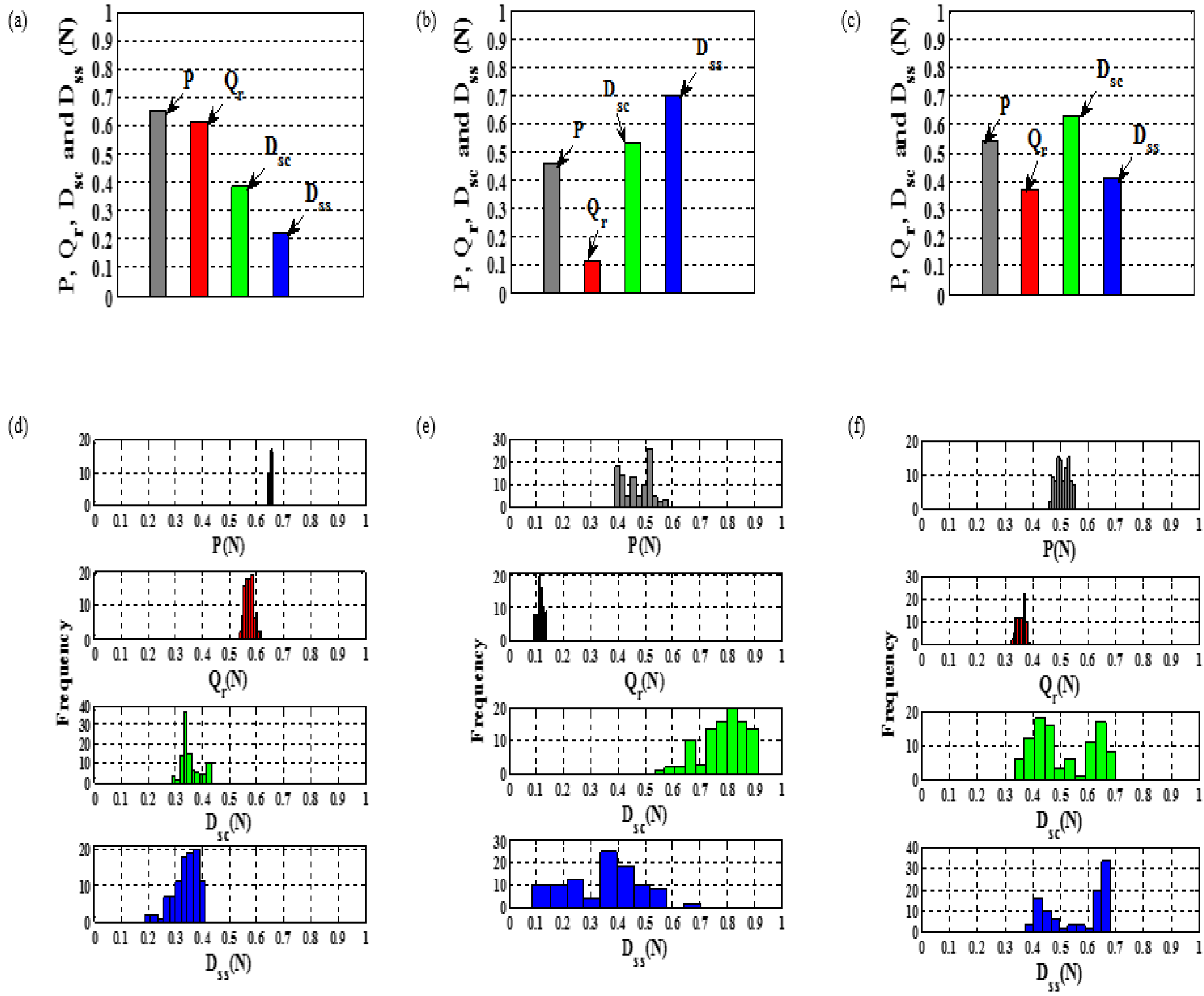


Figure 9. The normalised values of P, Q_r, D_{sc} and D_{ss} measured under sinusoidal supply voltage for (a) a dimmer controlled R-L impedance, (b) a number of computers and (c) a number of compact fluorescent lamps, and the histograms of the normalised powers measured under distorted test voltages for (d) a dimmer controlled R-L impedance, (e) a number of computers and (f) a number of compact fluorescent lamps.

On the other hand, it can be mentioned from Figure 8 (b), (c), (e) and (f) that the constant torque and constant speed cases of the induction machine have almost the same power values for sinusoidal and the distorted test voltages.

It is seen from Figure 9 (a) that P, Q_r, D_{sc} and D_{ss} of the R-L impedance controlled with a dimming circuit are 0.65, 0.61, 0.38 and 0.23 pu for sinusoidal supply voltage. Figure 9 (a) and (d) show that the P values drawn by the same load under sinusoidal and the distorted test voltages are very close. In addition to that, for the distorted test voltages, Q_r, D_{sc} and D_{ss} vary in the intervals from 0.5 to 0.6 pu, from 0.3 to 0.4 pu and from 0.2 to 0.4 pu, respectively.

It is observed from Figure 9 (b) that P, Q_r, D_{sc} and D_{ss} of the computers are 0.45, 0.11, 0.53 and 0.70 pu for sinusoidal supply voltage. Figure 9 (b) and (e) shows that the Q_r values drawn by the same load under sinusoidal and the distorted test voltages are very close. For the distorted test voltages, P, D_{sc} and D_{ss} vary in the intervals from 0.4 to 0.6 pu, from 0.5 to 0.9 pu and from 0.1 to 0.7 pu, respectively

It can be mentioned from Figure 9 (c) that the P, Q_r, D_{sc} and D_{ss} of the compact fluorescent lamps are measured as 0.54, 0.37, 0.63 and 0.41 pu in sinusoidal supply voltage case. For the distorted test voltage cases, the histograms plotted in Figure 9 (f) show that P is in the interval between 0.4 and 0.6 pu, Q_r is in the interval between 0.3 and 0.4 pu, D_{sc} is in the interval between 0.3 and 0.7 pu, and D_{ss} is in the interval between 0.3 and 0.7 pu.

From statistical results given above, one can see that D_{sc} could be successfully used to detect harmonic producing loads under sinusoidal and distorted supply voltages due to the fact that it has two distinct cases for the linear and nonlinear (harmonic producing) loads:

- The normalised values of D_{sc} measured for R-L impedance and induction machine are almost zero.
- However, for the harmonic producing loads, D_{sc} has large normalised values.

5. Conclusions

In this paper, a power resolution is proposed for unbalanced and nonsinusoidal systems. The motivation of the proposed resolution is to provide the direct determination of the power of optimum balanced capacitive compensator and to be used for detection of the harmonic producing loads in the smart power grids.

The simulation studies and analytical expressions show that the resolution achieves its compensation goal for the systems. In addition to that, the results demonstrated that it can practically be employed as a tool to design the cost effective unity power factor compensator consisting of the basic capacitors and an active compensator.

On the other hand, the scattered conductance power (D_{sc}) of the proposed resolution is statistically investigated for various load types and supply voltages in an experimental test system. Consequently, it is observed from the results that the normalised value of this power component is very close to zero for linear loads and it has considerably high value for nonlinear (harmonic producing) loads. Thus, it is pointed out that in the smart power grids harmonic producing loads could be detected by using the proposed resolution implemented in the demand meters.

Finally, due to fact that all power components of the proposed power resolution are related to the load conductance and susceptance parameters, it may provide a collective operation platform including not only the basic capacitors & active compensators, demonstrated as in this paper, but also other types of compensators. This will be studied in a future work.

6. References

[1] M. E. Balci, M. H. Hocaoglu, “A power resolution for nonsinusoidal and unbalanced systems — Part I: Literature overview and motivation”, *ELECO 2011*, Bursa, Turkey, 1-4 December 2011, pp. I-167 - I-172.

[2] M. E. Balci, A. E. Emanuel, “Apparent power definitions: A comparison study”, *IREE Journal*, vol. 6, no. 6, pp. 2713-2722, November 2011.

[3] M. E. Balci, M. H. Hocaoglu, “Quantitative comparison of power decompositions”, *Electric Power Systems Research*, vol. 78, no. 3, pp. 318-329, March 2008.

[4] M. E. Balci, M. H. Hocaoglu, “Comparative review of multi-phase apparent power definitions”, *ELECO 2009*, Bursa, Turkey, 5-8 November 2009, pp.I-144,I-148.

[5] M. E. Balci, M. H. Hocaoglu, “A current decomposition-based method for computationally efficient implementation of power resolution meters in non- sinusoidal single-phase systems”, *Metrology and Measurement Systems*, vol. 20, no. 2, pp. 263-274, June 2013.

[6] M. E. Balci, M. H. Hocaoglu, ‘‘New power decomposition for sinusoidal and nonsinusoidal conditions’’, *IEEE ICHQP 2006*, Cascais, Portugal, September 2006.

[7] M. E. Hocaoglu, M. H. Hocaoglu, “An analysis on the nonactive powers for nonsinusoidal conditions”, *Journal of the Faculty of Engineering and Architecture of Gazi University*, vol. 26, no. 2, pp. 307-313, June 2011 (in Turkish).

[8] M. E. Balci, M. H. Hocaoglu, “A power resolution for nonsinusoidal and unbalanced systems — Part II: Theoretical backgorund”, *ELECO 2011*, Bursa, Turkey, 1-4 December 2011, pp. I-173 - I-178.

[9] M. E. Balci, M. H. Hocaoglu, “Addendum to a power resolution for nonsinusoidal and unbalanced systems: Evaluation examples”, *ELECO 2011*, Bursa, Turkey, 1-4 December 2011, pp. I-179 - I-182.

[10] *AC Quantities- Part 1: Single-Phase Circuits, Part 2: Polyphase Circuits*, DIN Std. 40110, 1994 and 2002. (in German).

[11] *IEEE Standard Definitions for the Measurement of Electric Power Quantities under Sinusoidal, Non-sinusoidal, Balanced or Unbalanced Conditions*, IEEE Std. 1459, 2010.

[12] J. L. Willems, J. A. Ghijselen, A. E. Emanuel, ‘‘The apparent power concept and the IEEE Standard 1459-2000’’, *IEEE Trans. on Power Deliv.*, vol. 20, no. 2, pp. 876-884, April 2005.

[13] J. L. Willems, J. A. Ghijselen, A. E. Emanuel, ‘‘Addendum to the apparent power concept and the

IEEE Standard 1459-2000'', *IEEE Trans. on Power Deliv.*, vol. 20, no. 2, pp. 885-886, April 2005.
[14] L. S. Czarnecki, ''Physical reasons of currents RMS value increase in power systems with nonsinusoidal voltage'', *IEEE Trans. on Power. Deliv.*, vol. 8, no. 1, pp. 437-447, January 1993.
[15] L. S. Czarnecki, ''Orthogonal decomposition of the currents in a 3-phase nonlinear asymmetrical circuit with a nonsinusoidal voltage source'', *IEEE Trans. on Instrum. and Meas.*, vol. 37, no. 1, pp. 30-34, March 1988.
[16] H. Akagi, E. H. Watanabe, M. Aredes, "Instantaneous Power Theory and Applications to Power Conditioning", IEEE Press, John Wiley & Sons, 2007.
[17] *IEEE Recommended Practices and Requirements for Harmonic Control in Electrical Power Systems,* IEEE Std. 519, 2014.

7. Acknowledgement

This work was supported by Turkish Scientific Research Council (TUBITAK) under the project number 110E113.

Murat Erhan Balci received B.Sc. degree from Kocaeli University, M.Sc. and Ph.D. degrees from Gebze Institute of Technology, Turkey in 2001, 2004 and 2009, respectively. During 2008, he was a visiting scholar at Worcester Polytechnic Institute, USA. Since 2009, he has been with the Electrical and Electronics Engineering Department of Balikesir University, Turkey. He is, currently, an Associate Professor at the same University. He is working in the field of electric machines, power electronics, power quality, power system analysis and wind power.

Mehmet Hakan Hocaoglu received the B.Sc. and M.Sc. degrees from Marmara University, Turkey. He obtained the Ph.D. degree in 1999 from Cardiff School of Engineering, UK. From 1988 to 1993, he worked at Gaziantep University, Turkey as a Lecturer. Since 1999, he has been with the Electronics Engineering Department of Gebze Technical University, Turkey . He is, currently, a full Professor at the same University.

DESIGN OF WIRELESS POWER TRANSFER SYSTEM WITH TRIPLET COIL CONFIGURATION BASED ON MAGNETIC RESONANCE

Talal F. Skaik[1], Basel O. AlWadiya[2],

[1]Electrical Engineering Department, Islamic University of Gaza, P.O. Box 108, Palestine
[2]Engineering Department, Al-Azhar University – Gaza, P.O. Box 1277, Palestine
talalskaik@gmail.com, basmust@hotmail.com

Abstract: *Wireless power transfer (WPT) system based on magnetic resonance is presented here. The aim is to transfer energy wirelessly from transmitter coil to receiver coil based on magnetic resonance. A novel system with a two-coil transmitter connected to a single power source is proposed here in a triplet configuration with a single receiving coil. The two-coil transmitter is introduced as an extension to the converge area. The equivalent lumped element circuit model is presented and mathematical equations for scattering parameters have been derived. The proposed configuration is simulated using both circuit (ADS) and electromagnetic (EMPRO) simulators. The effect of the coupling between coils is investigated using simulation. The proposed configuration is practically implemented using solenoid coils and tested to verify the simulation results. The effect of receiver displacement on efficiency is also investigated.*

Keywords: *Magnetic resonance, Triplet configuration, Two-coil transmitter, Wireless Power Transfer.*

1. Introduction

In the recent years, an extensive research has been done on wireless power transfer, particularly between two coils across an air gap via magnetic coupling. Wireless power transfer can be used in many applications such as battery charging for portable electronic devices [1], electric vehicles [2,3], robots [4,5] and implantable medical devices [6,7]. In [8], researchers in Massachusetts Institute of Technology presented wireless energy transmission via strong magnetic resonant coupling with experimental demonstration. The transfer efficiency rapidly decreases as the transmission distance increases. Hence, several researches reported in literature are focused on increasing transfer distance as well as efficiency. To increase the transmission distance, extra coils called repeaters may be added between the transmitter and receiver as reported in [9]-[11]. Furthermore, the transfer efficiency is affected by the orientation and displacement of the receiving coil. In [12], the study demonstrates the efficiency dependence on receiver orientation and deviation.

Some applications need powering various devices at the same time and studies on multi-receiver wireless power transfer systems have been reported [13,14]. Other applications require multiple transmitters to cover larger area available for wireless power transfer, but most of the research reported in literature considered only a single transmitting coil. In [15], a parallel line feeder is proposed as the transmitter in wireless power transfer system to extend the coverage area. However, the efficiency of such system still needs further study for enhancement. In [16], a multiple-input multiple-output wireless power transfer systems is proposed whereby multiple separated transmission sources are used. However, this system still needs multiple voltage sources. Here we propose a wireless power transfer system using two-transmitting coils connected to a single voltage source so that the power of the single source is divided and delivered to both transmitting coils as depicted in Figure 1 (b). In such configuration, the coverage area is extended in comparison to conventional configuration in Figure 1 (a). Mathematical equations have been derived for the proposed system and circuit simulation has been performed based on the equivalent lumped element model. Moreover, Electromagnetic professional program EMPRO [17] has been utilized to model the proposed configurations. The proposed system is practically implemented and tested to verify the simulation results.

2. Equivalent Lumped Element Model

The lumped element equivalent circuits for the single-coil transmitter and two-coil transmitter systems are shown in Figure 2 (a) and Figure 2 (b), respectively. The AC voltage source has internal resistance R_s and the driving loop is represented by internal resistance R_1 and self

Received on: 24.11.2016
Accepted on: 08.02.2017

inductance L_1. Similarly, the load loop is represented by internal resistance and self inductance connected to a load Z_L=R_L.

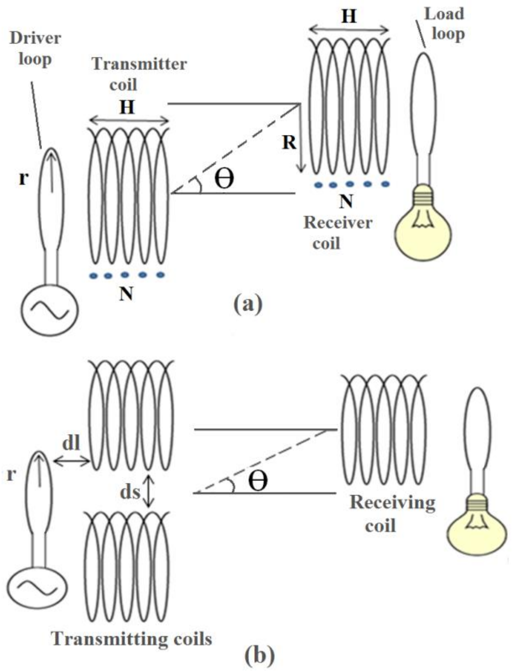


Figure 1. Wireless power transfer system (a) Single-coil transmitter configuration, (b) Proposed two-coil transmitter configuration

The equivalent model of a transmitting or receiving coil is a series *RLC* resonating circuit with an internal resistance *R*, self inductance *L* and capacitance *C*. The mutual coupling between coils is represented in the model by inductive coupling coefficient $K_{i,j}$ between coil *i* and coil *j* and it is calculated by:

$$K_{i,j} = M_{i,j} \big/ \sqrt{L_i L_j} \qquad (1)$$

where $M_{i,j}$ is the mutual inductance between coils *i* and *j* and L_i and L_j are self-inductances of coils *i* and *j* respectively. The self-inductance of coil *i* is calculated by [9],

$$L_i = 4\pi \times 10^{-7} \times n_i^2 \times r_i \left[ln\left(\frac{8r_i}{a_i}\right) - 2 \right] \qquad (2)$$

where n_i and r_i are the number of turns and radius of coil *i* respectively, and a_i is the radius of copper wire for coil *i*. The mutual inductance $M_{i,j}$ between coils *i* and *j* is calculated by [12],

$$M_{i,j} = \frac{\pi \times \left(4\pi \times 10^{-7}\right) \times \sqrt{n_i n_j} \times r_i r_j}{2D^3} \qquad (3)$$

where n_i and n_j are number of turns of coils *i* and *j*, r_i and r_j are the radii of coils *i* and *j* and *D* is the distance between the two coils.

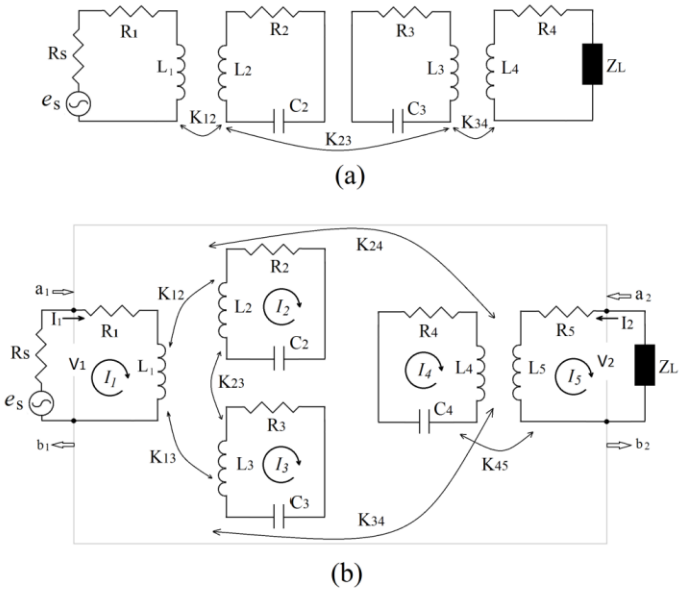


Figure 2. Equivalent circuit models for configurations in Figure 1 (a) Model for system in Figure 1 (a), (b) Model for system in Figure 1 (b).

The operating angular frequency for the single-coil transmitter in Figure 2 (a) is $\omega_o=1/(L_2C_2)^{0.5}=1/(L_3C_3)^{0.5}$ and for the two-coil transmitter in Figure 2 (b) is $\omega_o=1/(L_2C_2)^{0.5}=1/(L_3C_3)^{0.5} = 1/(L_4C_4)^{0.5}$. The mathematical equations for the conventional model in Figure 2 (a) are reported in literature and the mathematical description of the proposed model in Figure 2 (b) can be written in matrix form as $[Z][i]=[e]$ where

$$[i] = \begin{bmatrix} i_1 \\ i_2 \\ i_3 \\ i_4 \\ i_5 \end{bmatrix} \quad , \quad and \quad [e] = \begin{bmatrix} e_s \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad (4)$$

$$[Z] = \begin{bmatrix} Z_{11} & j\omega M_{12} & j\omega M_{13} & 0 & 0 \\ j\omega M_{21} & Z_{22} & j\omega M_{23} & j\omega M_{24} & 0 \\ j\omega M_{31} & j\omega M_{32} & Z_{33} & j\omega M_{34} & 0 \\ 0 & j\omega M_{42} & j\omega M_{43} & Z_{44} & j\omega M_{45} \\ 0 & 0 & 0 & j\omega M_{54} & Z_{55} \end{bmatrix} \qquad (5)$$

where $[Z]$ is the impedance matrix and

$$\begin{aligned} Z_{11} &= R_1 + R_s + j\omega L_1 \\ Z_{22} &= R_2 + j\omega L_2 + \left(1/j\omega C_2\right) \\ Z_{33} &= R_3 + j\omega L_3 + \left(1/j\omega C_3\right) \\ Z_{44} &= R_4 + j\omega L_4 + \left(1/j\omega C_4\right) \\ Z_{55} &= R_5 + j\omega L_5 + R_L \end{aligned} \qquad (6)$$

The $[Z]$ matrix can be written in term of coupling coefficients K_{ij} using (1) by replacing $j\omega M_{ij}$ by

$\omega K_{ij}\sqrt{L_i L_j}$. The scattering parameters of the circuit are found as,

$$S_{11} = \left.\frac{b_1}{a_1}\right|_{a_2=0} \quad , \quad S_{21} = \left.\frac{b_2}{a_1}\right|_{a_2=0} \tag{7}$$

where b_1, b_2 represent waves reflected from ports 1 and 2, respectively, and a_1, a_2 represent waves incident on ports 1 and 2 respectively, and they can be found by:

$$a_N = \frac{1}{2}\left(\frac{V_N}{\sqrt{R}} + \sqrt{R} I_N\right)$$
$$b_N = \frac{1}{2}\left(\frac{V_N}{\sqrt{R}} - \sqrt{R} I_N\right) \tag{8}$$

where V_N and I_N represent the voltage and current at port N and they are found from the equivalent circuit in Figure 2 (*b*) as $V_1=e_s$ - I_1R_S and $V_2=-I_2R_L= I_5R_L$. By substitution of V_1, V_2, I_1 and I_2 into (8) the wave equations are now found in terms of circuit parameters as follows:

$$a_1 = \frac{e_S}{2\sqrt{R_s}}, \quad b_1 = \frac{e_S - 2I_1R_s}{2\sqrt{R_s}}, \quad b_2 = I_5\sqrt{R_L} \tag{9}$$

The scattering parameters of the circuit are now found in terms of circuit parameters by substitution of (9) into (7) as follows:

$$S_{11} = 1 - \frac{2I_1R_s}{e_s} \quad and \quad S_{21} = \frac{2\sqrt{R_sR_L}\,I_5}{e_s} \tag{10}$$

Solving [*Z*][*i*]=[*e*] for currents I_1 and I_5 we obtain,

$$I_1 = [Z]^{-1}_{11} e_s \quad and \quad I_5 = [Z]^{-1}_{51} e_s \tag{11}$$

and by substitution of (11) into (10), the scattering parameters are now found in terms of impedance matrix as follows,

$$S_{11} = 1 - 2R_s[Z]^{-1}_{11} \quad , \quad S_{21} = 2\sqrt{R_sR_L}[Z]^{-1}_{51} \tag{12}$$

3. Two-coil transmitter system simulation

The configuration of single transmitting coil is extensively studied and presented in literature. To extend the coverage area, the two-coil transmitter configuration in Figure 1 (b) is investigated here. Here we use a single driving loop connected to a single AC voltage source. The power is delivered from the driving loop and is divided to both the transmitting coils. The two-coil transmitter configuration is simulated firstly using ADS circuit simulator and then using EMPRO electromagnetic simulator.

3.1. Circuit Simulation

The system equivalent lumped-element circuit simulated using ADS is depicted in Figure 3 whereby coupling is represented by coupling coefficient *K*. The internal resistance of any coil is assumed 0.5 Ω and the transmitting and receiving coils have capacitance *C*=5 pF and inductance *L*=75 μH which give resonant frequency of 8.21873 MHz. The single-turn driving and receiving loops have an inductance of 1.7387 μH and R_s=R_L=50 Ω. The effect of the couplings between coils on power transfer efficiency is investigated by observing the scattering parameters S_{21} and S_{11}. The efficiency (*η*) is calculated directly by $\eta=|S_{21}|^2$ x 100%.

The coupling between the transmitting coils K_{23} is firstly investigated to study the effect on efficiency. Other coupling coefficients are kept constant with K_{12}=K_{13}=K_{45}=0.2, and K_{24}=K_{34}=0.04. The coupling K_{23} is changed from 0.07 to 0.28 and the S_{21} magnitude response is observed as depicted in Figure 4. It can be noticed that the transmission efficiency degrades as coupling between the transmitting coils is increased.

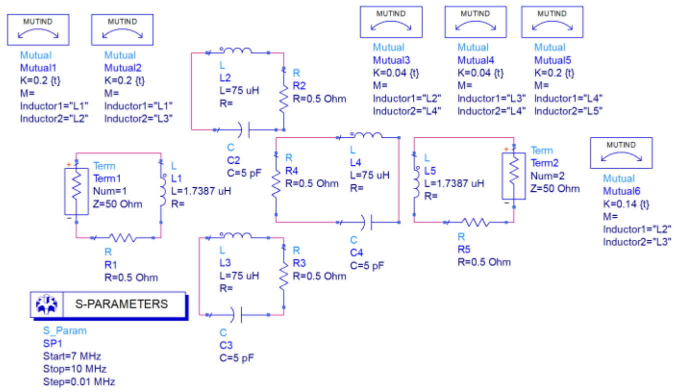


Figure 3. Two-coil transmitter ADS circuit.

Furthermore, the coupling between the transmitting coils and the receiving coil (K_{24} and K_{34}) is also investigated by varying the coupling K_{24} and K_{34} from 0.01 to 0.05 while keeping other coupling coefficients fixed as K_{12}=K_{13}=K_{45}=0.2 and K_{23}=0.14. The S_{21} magnitude response is shown in Figure 5 and it can be noticed that as the coupling decreases the efficiency decreases.

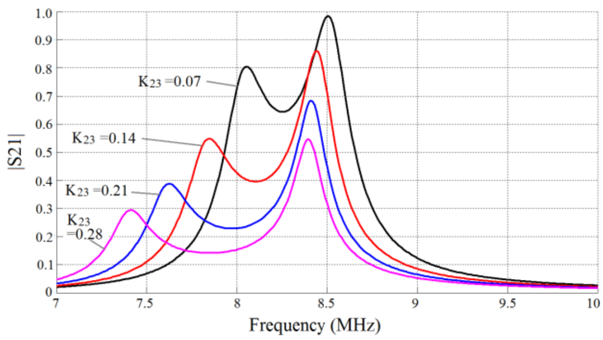


Figure 4. S_{21} magnitude response for various values of coupling K_{23} using ADS.

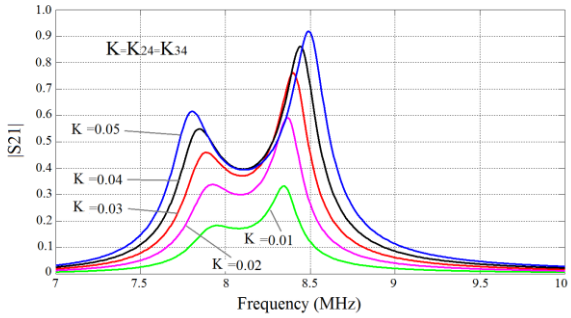


Figure 5. S_{21} magnitude response for various values of coupling K_{24}= K_{34}.

3.2. Electromagnetic Simulation

The two-coil transmitter configuration is simulated using electromagnetic simulator EMPRO to model the physical structure of the proposed configuration in Figure 1 (b). The transmitting and receiving coils are identical solenoid coils with height (*H*), radius (*R*), and number of turns (*N*). The transmission coils and the receiving coil are separated by a distance of 1 meter. The system parameters of the coils are set with fixed values: *H*=16 Cm, *N*= 5.5 and *R*=35 Cm. Other parameters such as the distance between two transmitters (*ds*), the distance between the transmitters and the driving loop (*dl*), and the radius of the driving loop (*r*) are set as variables. A parametric study has been carried out on these parameters and best efficiency of 66.49% is obtained at *r*=290 mm, *dl*=65 mm and *ds*= 30 mm. The simulated scattering parameters are shown in Figure 6 and the best transfer efficiency is obtained at frequency 8.4072 MHz. The structure of the two-coil transmitter system along with the magnetic field distribution is shown in Figure 7. The axis of the receiving coil lies in the middle of separation distance between the two transmitting coils.

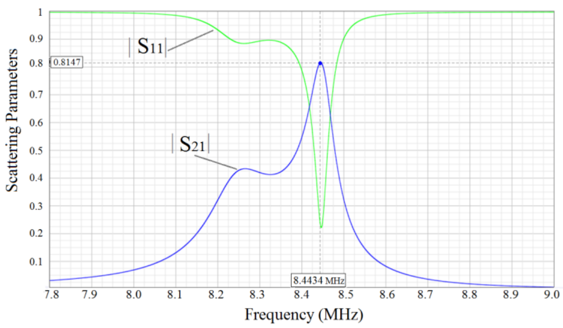


Figure 6. Simulated S-parameters for two-coil transmitter configuration**.**

4. Simulation of receiver displacement

The angle θ representing the displacement between the transmitter and receiver coils is also investigated to see its effect on the efficiency by using EMPRO. Table 1 shows the efficiency versus θ for the proposed configuration. It is noticed that the highest efficiency is obtained when (θ=0) that is when the axis of the receiving coil lies in the middle of separation distance between the two transmitting coils. Moreover, it is clear that when the angle θ is increased by displacing the receiving coil away from concentric axis the efficiency decreases. As we increase the angle θ from 0º to 35º, the efficiency for two-coil transmitter system drops from 66.49% to 40.37%.

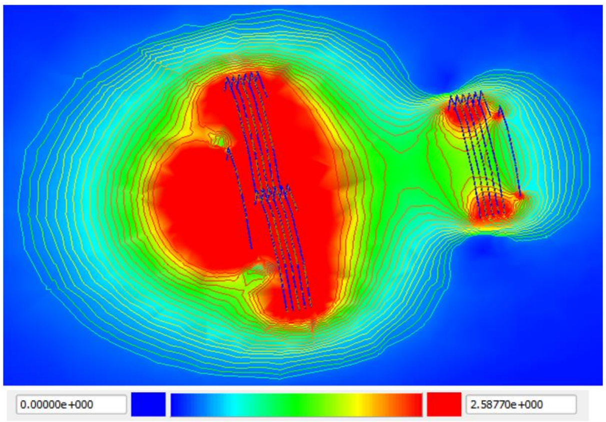


Figure 7. Magnetic field plots for two-coil transmitter configuration

Table 1. Efficiency versus receiver displacement

Angle (θº)	Efficiency (η) for two coil transmitter
0	66.49%
5	63.99%
10	61.45%
15	58.03%
20	57.65%
25	53.75%
30	44.35%
35	40.37%

5. Experimental results

To verify the simulation results, the proposed two-coil transmitter configuration has been implemented practically. The coils are made of copper wires of the same length and wire diameter of 2.75 mm and they are wounded around wooden cylinders to from solenoid coils. The parameters of the coils are those obtained from simulation results and they are *N*=5.5, *H*=160 mm and *R*=350 mm. A signal generator with frequency up to 10 MHz is used as the source for the transmitter and a digital oscilloscope has been utilized to show the received signal.

The two-coil transmitter experimental model is shown in Figure 8. The distance between the transmitting and receiving coils is 1 meter and the parameters of the system are *ds*= 30 mm, *dl*= 65 mm, and *r*= 290 mm. A sinewave of voltage amplitude 10 V is set at the generator and its frequency is tuned until maxim power transfer is obtained at resonant frequency of 9.3 MHz. The voltage at the receiving coil terminals was measured and the efficiency was then calculated and found about 61.01%.

The efficiency (η) has been experimentally found for different receiver displacement angles (θ) and the results are shown in Table 2. It can be noticed that the simulation results in Table 1 are slightly different from the experimental results due to imperfection in fabrication of coils that resulted in some changes in the dimensions in the structure. Moreover, the ohmic losses of the conducting coils contributed in efficiency degradation.

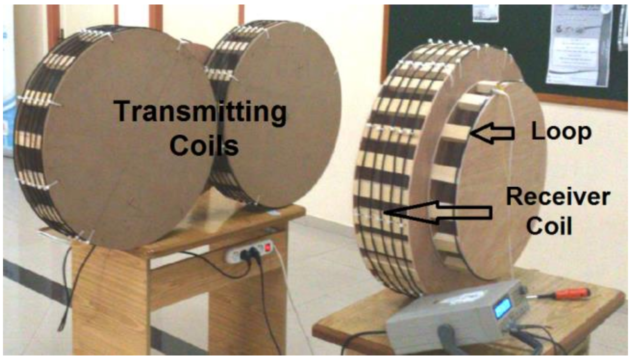


Figure 8. Experimental model for two-coil transmitter configuration

It is noticed that the highest simulated efficiency is 66.49% while the experimental efficiency is 61.01% that is about 5% lower than the simulated. The experimental results show a drop of efficiency of 30% when the angle θ is increased from 0^o to 35^o.

Table 2. Experimental results of efficiency versus receiver dispalcement

Angle (θ^o)	Two-Coil Transmitter Experimental (η)
0	61.01%
5	57.32%
10	48.10%
15	44.23%
20	40.61%
25	38.25%
30	32.56%
35	30.90%

6. Conclusions

In this paper, a two-coil transmitter single-coil receiver wireless power transfer system is presented. The system is proposed as an extension for the coverage area in comparison to conventional single-transmitter configurations. Mathematical model has been derived and both circuit and electromagnetic simulations have been done to show the scattering parameters and transfer efficiency. Receiver displacement has also been investigated and experimental solenoid coil models have been implemented.

6. References

[1] S.Y. Hui, "Planar wireless charging technology for portable electronic products and Qi," *in Proceedings of the IEEE* , vol.101, no.6, pp.1290-1301, June 2013.

[2] C-S. Wang, O.H. Stielau, and G.A.Covic, "Design considerations for a contactless electric vehicle battery charger," *in IEEE Transactions on Industrial Electronics*, vol.52, no.5, pp.1308-1314, Oct. 2005.

[3] T. Imura, H.Okabe, and Y.Hori, "Basic experimental study on helical antennas of wireless power transfer for Electric Vehicles by using magnetic resonant couplings," in *proceedings of Vehicle Power and Propulsion Conference*, pp.936-940, 7-10 Sept. 2009.

[4] I-K. Cho, S-M Kim, J-I. Moon, J-H. Yoon; S-I. Jeon, and J-I. Choi, "Wireless power transfer system for docent robot by using magnetic resonant coils," in *proceedings of IEEE 5th International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications (MAPE)*, pp.251-254, 29-31 Oct. 2013.

[5] T. Deyle and M. Reynolds, "Surface based wireless power transmission and bidirectional communication for autonomous robot swarms," in *proceedings of IEEE International Conference on Robotics and Automation* (ICRA), pp.1036-1041, 19-23 May 2008.

[6] X. Li; C-Y. Tsui, and W-H. Ki, "A 13.56 MHz wireless power transfer system with reconfigurable resonant regulating rectifier and wireless power control for implantable medical devices," in *IEEE Journal of Solid-State Circuits*, vol.50, no.4, pp.978-989, April 2015.

[7] J. Walk, J. Weber, C. Soell, R. Weigel, G. Fischer, and T. Ussmueller, "Remote powered medical implants for telemonitoring," in *Proceedings of the IEEE* , vol.102, no.11, pp.1811-1832, Nov. 2014.

[8] A. Kurs, A. Karalis, R. Moffatt, J. D. Joannopoulos, P. Fisher, and M. Soljacic, "Wireless power transfer via strongly coupled magnetic resonances," *Science*, vol. 317, no. 6, pp. 83–86, 2007.

[9] B. Luo, S. Wu, and N. Zhou, "Flexible design method for multi-repeater wireless power transfer system based on coupled resonator bandpass filter model," *in IEEE Transactions on Circuits and Systems*, vol.61, no.11, pp.3288-3297, Nov. 2014.

[10] T. Imura, "Equivalent circuit for repeater antenna for wireless power transfer via magnetic resonant coupling considering signed coupling," *in proceeding of 6th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp.1501-1506, 21-23 June 2011.

[11] D. Ahn and S. Hong, "A study on magnetic field repeater in wireless power transfer," *in IEEE Transactions on Industrial Electronics*, vol.60, no.1, pp.360-371, Jan. 2013.

[12] Z. Yan, Y. Li, C. Zhang, and Q. Yang, "Influence factors analysis and improvement method on efficiency of wireless power transfer via coupled magnetic resonance," in *IEEE Transactions on Magnetics*, vol.50, no.4, pp.1-4, April 2014.

[13] V. Jiwariyavej, T. Imura, and Y. Hori, "Coupling coefficients estimation of wireless power transfer system via magnetic resonance coupling using information from either side of the system," in *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol.3, no.1, pp.191-200, March 2015.

[14] K.E. Koh, T.C. Beh, T. Imura, and Y. Hori, "Multi-receiver and repeater wireless power transfer via magnetic resonance coupling - Impedance matching and power division utilizing impedance inverter," in *proceeding of 15th International Conference on Electrical Machines and Systems (ICEMS)*, pp.1-6, 21-24 Oct. 2012.

[15] T. Higashino, M. Ziji, M. Okada, Y. Tatsuta, Y. Goto, Y. Tsuruda, and R. Tanaka, "A new configuration of magnetic coupled power transfer using parallel line feeder," in *proceedings of Wireless Power Transfer (WPT) conference*, pp.171-174, 15-16 May 2013.

[16] M.Q. Nguyen, Y. Chou, D. Plesa, S. Rao, and J-C. Chiao, "Multiple-inputs and multiple-outputs wireless power combining and delivering systems," in *IEEE Transactions on Power Electronics*, vol.30, no.11, pp.6254-6263, Nov. 2015.

[17] http://www.keysight.com/en/pc-1297143/empro-3d-em-simulation-software?&cc=PS&lc=eng

Talal Skaik received the B.Sc. degree in 2004 from the Islamic University of Gaza, where he worked as a teaching assistant until 2006. He was awarded Hani Qaddumi scholarship and received M.Sc. degree in Communications Engineering with distinction in 2007 from the University of Birmingham, UK. He was awarded the ORSAS scholarship for doctoral study at the University of Birmingham, UK and received PhD degree in Microwave Engineering in 2011. Throughout his PhD study, he worked as a teaching assistant, and also as a research associate on micromachined microwave circuits. He was the Head of Electrical Engineering Department at the Islamic University of Gaza from Sept. 2014 until August 2016. He is currently an assistant professor at the Islamic University of Gaza. His research interests include design of microwave filters, diplexers, multiplexers, energy harvesting systems, reconfigurable antennas and microwave passive components.

Basel AlWadiya received BSc degree from Misr University for science and technology and MSc degree in communications engineering from the Islamic University of Gaza in 2012. He is currently working at AlAzhar University in Gaza.

Yener TASKIN / IU-JEEE Vol. 17(1), (2017), 3175-3180

FUZZY PID CONTROLLER FOR PROPELLER PENDULUM

Yener TASKIN

Department of Mechanical Engineering, Istanbul University, Turkey
ytaskin@istanbul.edu.tr

Abstract: *In this paper, a fuzzy PID controller is proposed for angular position control of a nonlinear propeller pendulum system. While classical control methods work well on linear systems, nonlinear control approaches should be designed for nonlinear ones. On the one hand, there are three constant gains related with linear proportional, integral and derivative terms in classical PID controller. On the other hand, these gains are varied with time by the proposed controller using fuzzy logic inference. In order to demonstrate the position control enhancement for the nonlinear system, the proposed controller is compared with classical PID controller using simulation results with and without external disturbance. The simulation results show that the proposed Fuzzy PID controller is more successful in reference tracking than classical PID controller.*

Keywords: *Fuzzy logic, PID, angular position, nonlinear control, propeller pendulum***.**

1. Introduction

Propeller pendulum is a type of compound propeller which is described as having a motorized propeller producing a thrust force at the end of a pendulum rod that can lift the pendulum up and down [1]. The thrust force can be utilized to stabilize the pendulum at any desired position using various control methods [1-7]. A PID method is preferred to control a driven pendulum in [2]. Kizmaz et al. also proposed a sliding mode controller using the linear part of the mathematical model of suspended pendulum system [3].

Propeller pendulum is assumed to be a simplified plant model of an unmanned autonomous vehicle which is used for teaching system dynamics and control topics in mechanical and mechatronics engineering education [8-11]. Huba et al. utilized a pendulum control to demonstrate different aspects of robust and nonlinear control for “learning by playing”, “learning by discovering”, or through “experiential learning” approaches during engineering education [8].

Propeller pendulum is a nonlinear system and can be controlled well by classical control methods when the system is linearized around design point. But, if the controlled state goes far away from the design point and/or the system is highly nonlinear, at this point, classical control methods such as PID control can hardly hold the dynamic performance [10]. Therefore, nonlinear pendulum system is a significant candidate to develope a nonlinear controller. In literature, various structures for fuzzy PID controllers have been proposed for various applications since fuzzy controllers demonstrate successful results [12-14]. In this paper, a fuzzy PID controller with a new structure is proposed for a nonlinear propeller pendulum system. Primarily, propeller pendulum is presented in the next section. Afterwards, the proposed Fuzzy PID controller is described briefly and then the discussion of simulation results and conlusions are given in the following sections respectively.

2. Propeller Pendulum

Figure 1 exhibits propeller pendulum model with an external disturbance force F_d applied perpendicular to pendulum rod. In this model, m_1 and m_2 represent the mass of pendulum rod and motorized propeller respectively. Desired angular position *theta* is measured during propeller thrust force T is applied. L is the length of propeller pendulum. In equation (1), nonlinear mathematical model of propeller pendulum is given. J and c are the moment of inertia of propeller pendulum and viscous damping coefficient respectively. g is the acceleration of gravity. Model parameters and their values are given in Table 1 [11].

Table 1. Model parameters [11]

Parameter	*SI Units*	*Value*
m_1	kg	0.21
m_2	kg	0.16
L	m	0.6
J	$\text{kg}\,\text{m}^2$	0.083
c	$\text{kg}\,\text{m}^2/s$	0.074
g	m/s^2	9.81

Received on: 15.08.2016
Accepted on: 19.12.2016

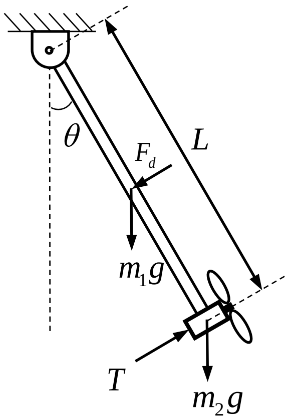


Figure 1. Propeller pendulum model

$$J\ddot{\theta}+c\dot{\theta}+\left(\frac{m_1}{2}+m_2\right)gL\sin\theta=TL-F_d\frac{L}{2} \quad (1)$$

3. Fuzzy PID Controller

In classical PID control, the present, past and future errors are compensated by linear proportional, integral and derivative terms. Thrust force can be obtained by the sum of these terms as seen in equation (2) where K_P, K_I and K_D are constants. Error is defined by difference between desired reference angular position and angular position of pendulum given in equation (3).

$$T_{PID}=K_Pe+K_I\int edt+K_D\frac{de}{dt} \quad (2)$$

$$e=\theta_{ref}-\theta \quad (3)$$

In proposed fuzzy PID control, the gains are not constant anymore. Each controller gain is calculated by a fuzzy logic unit and varied with time. This unit is shown in Figure 2 with a single input - single output relation.

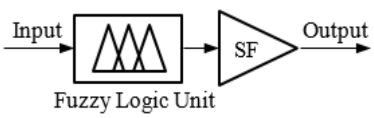


Figure 2. Fuzzy logic input-output representation

Time varied gains of Fuzzy PID are obtained from related fuzzy logic unit. The input is the related variable and the output is the related variable gain of the proposed controller. SF is the related scaling factor. In Table 2, the input-output relations are given for the related terms of proposed fuzzy PID controller. If input of the fuzzy logic unit is chosen as error (e), then scaling factor and output of the unit are taken as proportional scaling factor (PSF) and gain of the proportional controller (K_{FP}). If input of the fuzzy logic unit is chosen as integral of error ($\int edt$), then scaling factor and output are taken as integral scaling factor (ISF) and gain of the integral controller (K_{FI}). If input of the fuzzy logic unit is chosen as derivative of error (de/dt), then scaling factor and output are taken as derivative scaling factor (DSF) and gain of the derivative controller (K_{FD}). Thrust force generated from the proposed fuzzy PID controller is the sum of these terms and is obtained by equation (4).

$$T_{FuzzyPID}=K_{FP}e+K_{FI}\int edt+K_{FD}\frac{de}{dt} \quad (4)$$

The structure of the proposed controller is given in Figure 3. Block diagram shown in Figure 4 exhibits the control thrust force applied to propeller pendulum in the closed loop form.

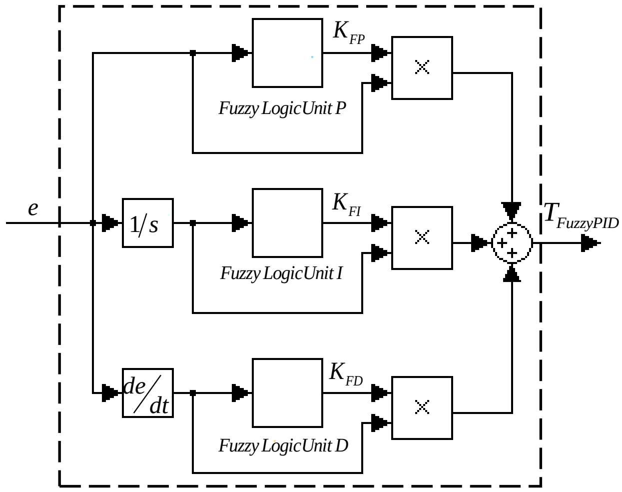


Figure 3. Structure of Fuzzy PID controller

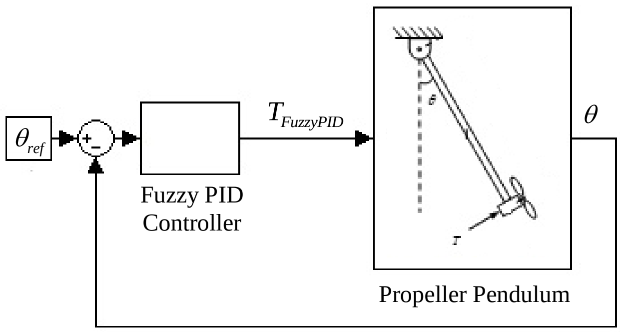
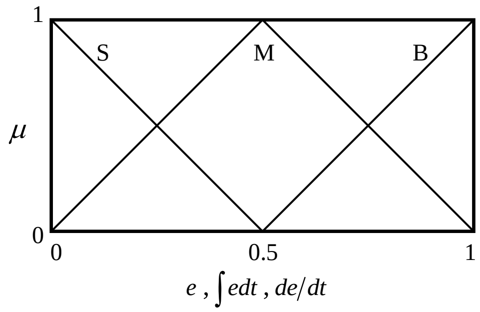


Figure 4. Block diagram of closed loop fuzzy PID controller of propeller pendulum

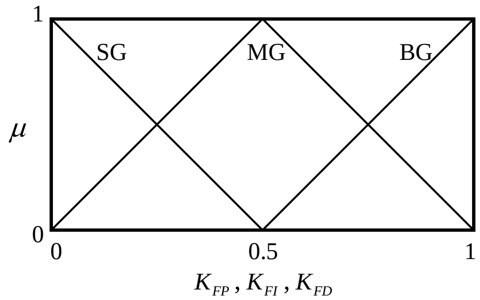
Table 2. Fuzzy logic input-output relation

Input	*Scaling Factor (SF)*	*Output*
e	PSF	K_{FP}
$\int edt$	ISF	K_{FI}
de/dt	DSF	K_{FD}

For each fuzzy logic unit of the proposed controller, Mandani type fuzzy inference with triangular membership functions is utilized and centroid method is used for defuzzification. Membership functions are shown in Figure 5 for input and output variables.



(a)



(b)

Figure 5. Membership functions a) input variables b) output variables

Fuzzy rule base is very simple and given in Table 3. It involves the same rules for each fuzzy logic unit.

Table 3. Rule base of each fuzzy logic unit

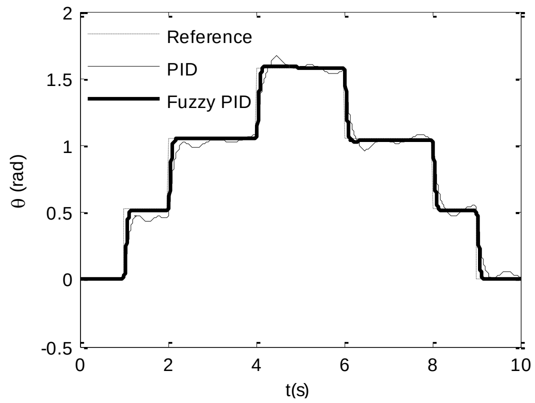
Input			***Output***		
e	$\int edt$	de/dt	K_{FP}	K_{FI}	K_{FD}
S			SG		
M			MG		
B			BG		

- If input is small (S) then output gain is small (SG)
- If input is medium (M) then output gain is medium (MG)
- If input is big (B) then output gain is big (BG)

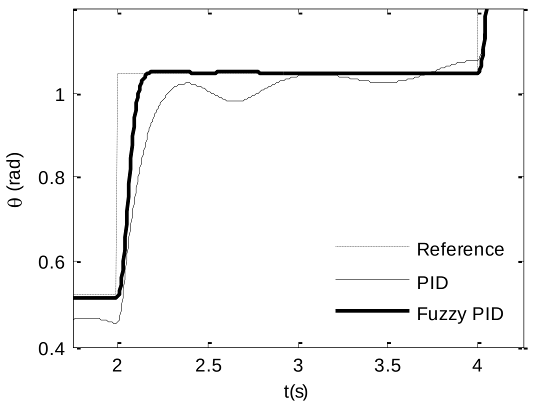
For instance, if angular position error is small, then proportional gain will be small. On the other hand, if the error is big, then the proportional gain will be big. Scaling factors of related terms are tuned by trial and error during simulation. Proportional (PSF), integral (ISF) and derivative (DSF) scaling factors are taken as 250, 100 and 11 respectively. Constant gains of classical PID controller are taken as 150, 50 and 2 for proportional (K_P), integral (K_I) and derivative (K_D) terms respectively. Thrust force is limited to ±20 N for motorized propeller.

4. Simulation Results

Position control of the propeller pendulum is evaluated by simulation results with and without external disturbance cases. Simulation duration and sampling time is taken as 10 and 0.005 seconds, respectively. Equation of motion is solved by Runge-Kutta method. Primarily without external disturbance simulation is made. Position reference is described by ascending and descending steps shown in Figure 6 a. The simulated position of the propeller pendulum is also demonstrated in the same figure for both classical PID and proposed fuzzy PID controllers. Pendulum stays still at the beginning position for one second. Afterwards, the position reference ascends to $\pi/6$ rad, $\pi/3$ rad and $\pi/2$ rad steps respectively. Then, pendulum follows the reference by descending to the beginning position symmetrically. While pendulum follows the reference very close in fuzzy PID controlled case, PID controlled case demonstrates different responses for each step during ascending and descending. PID controller demonstrates weak tracking performance since classical linear control methods do not work well on controlling nonlinear systems. This situation can be seen in Figure 6 b clearly. On the other hand, Fuzzy PID controlled pendulum follows the reference track without overshooting in each ascending and descending steps.



(a)



(b)

Figure 6. Angular position of propeller pendulum a) for 10 s duration b) between 1.75 s and 4.25 s

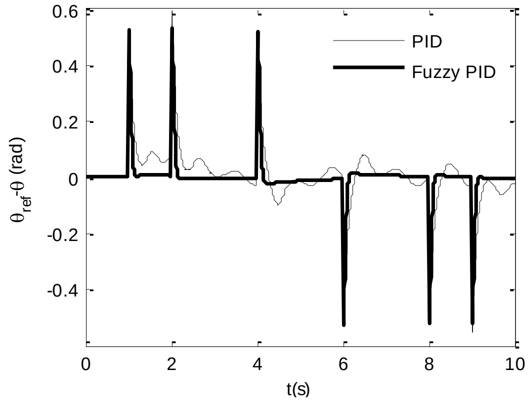


Figure 7. Error of angular position

Reference tracking success of the proposed controller can be evaluated in angular positon error diagram as well as angular position diagram. In Figure 7, it is easy to see that the angular position error of fuzzy PID controller goes to zero rapidly after each step reached up and down. On the other hand, in classical PID case, the position of pendulum can’t reach to step value or remains positive or negative position errors in each step.

While the gains are constant in classical PID controller, the gains of fuzzy PID controller are varied with time according to fuzzy inference which is described in the previous section. Variation of each control gain of fuzzy PID is plotted in Figure 8 for reference tracking simulation shown in Figure 6 a. Fuzzy proportional (K_{FP}), integral (K_{FI}) and derivative (K_{FD}) gains are changed as angular position error, integral of angular position error and derivative of angular position error are changed with the nonlinear behavior of the pendulum system. For instance, when the error increases, proposed fuzzy PID controller generates greater proportional gain in order to achieve better tracking performance. On the other hand, when the error goes to zero, it descends to lower values. The variation of the other two gains indicates the same character as the integral of error or the derivative of error increases, the related gains take greater values.

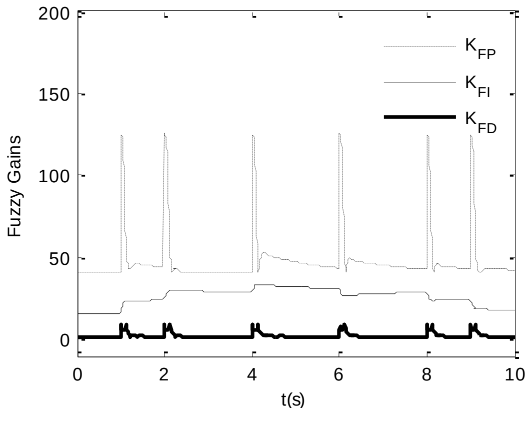


Figure 8. Variation of fuzzy PID gains

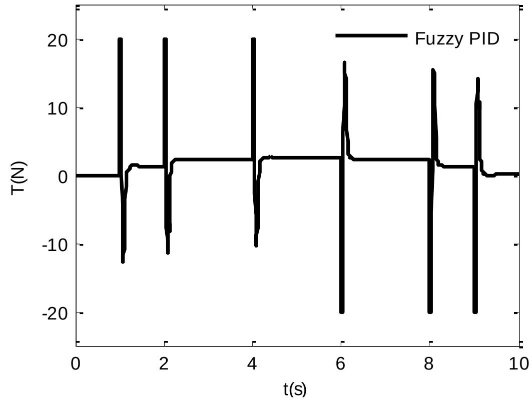


Figure 9. Thrust force

Thrust force applied by fuzzy PID controller is shown in Figure 9. The thrust force is saturated at the very beginning of ascending and descending positions during maximum position errors occur. When the controller generates a thruster force greater than the saturation value of the motorized propeller, it is limited to ±20 N. The thrust force reaches to a steady state value that holds the pendulum at the desired position as the position error goes to zero. In Figure 10, the external disturbance force acted on propeller pendulum is shown. Sine function is utilized to form external disturbance whose amplitude and frequency are 0.5 N and 1 Hz at 5 N bias.

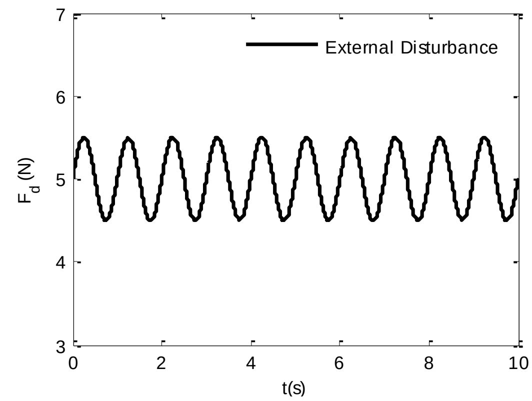


Figure 10. External disturbance

At this point, simulation results for the external disturbance applied case are given for the same reference tracked in the previous simulation in order to evaluate the performance of control under external disturbances. In the first one second duration, pendulum moves to negative direction under the effect of external disturbance in Figure 11 a. It is seen that the angular position of the propeller for the Fuzzy PID controlled case is much closer to the reference that is followed when it is compared with the classical PID controlled case in Figure 11 a and b.

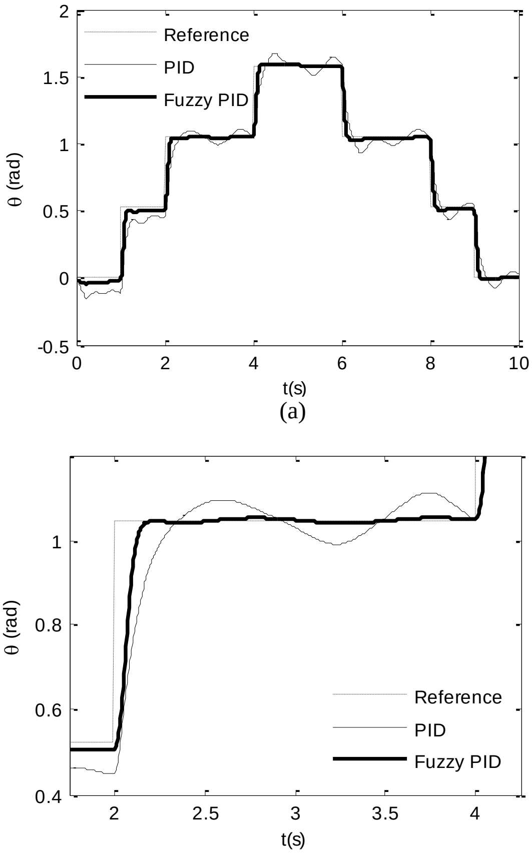


Figure 11. Angular position of propeller pendulum a) for 10 s duration b) between 1.75 s and 4.25 s

If the angular position of the pendulum is compared with the previous simulation given in Figure 6 a and b, it is easy to spot the track performance loss of the classical PID controlled case which is much worse by the disturbance. On the other hand, the Fuzzy PID controlled case still follows the reference as close as the previous simulation that can also be observed from the error of angular position diagram shown in Figure 12.

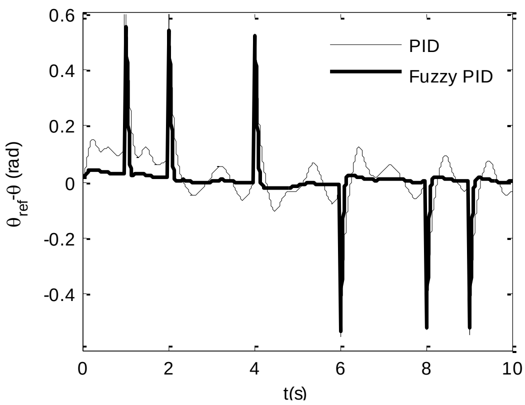


Figure 12. Error of angular position

As the pendulum moves to the negative direction at the beginning of the simulation under the effect of external disturbance, the proposed controller gives a sudden response pulling the pendulum as close as to the reference that can be seen in Figures 11 a and 14. Also, the change of fuzzy proportional gain in the first one second duration is observed in Figure 13 when compared with Figure 8.

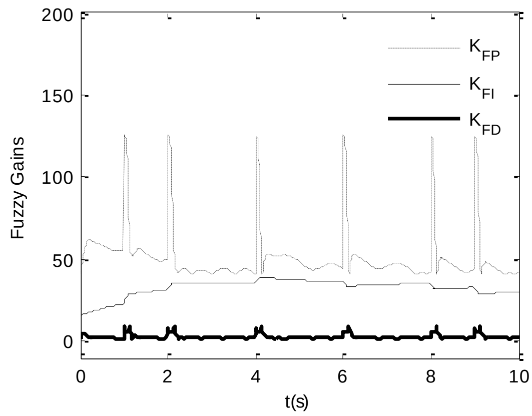


Figure 13. Variation of fuzzy PID gains

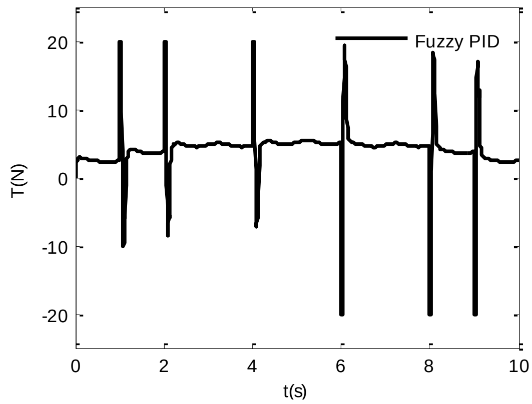


Figure 14. Thrust force

The reference tracking success of the proposed controller can also be determined from the root mean square (RMS) error comparisons given in Table 4. Case 1 and 2 represent without and with external disturbance cases respectively. In all cases, total RMS errors of fuzzy PID controller is less than PID controller ones. Total RMS thrust force values of all cases are also given in the same table. Although the thrust force need for fuzzy PID controller is greater than PID controller in all cases, the change in RMS error of fuzzy PID controller is less comparing to PID controller ones.

Table 4. RMS error and RMS thrust force comparisons

	Case 1		Case 2	
	PID	Fuzzy PID	PID	Fuzzy PID
e_{RMS}	0.115	0.091	0.130	0.094
T_{RMS}	3.275	4.473	5.028	5.882

5. Conclusions

A Fuzzy PID controller is designed and applied to a nonlinear propeller pendulum. The proposed controller which compensates the nonlinear character of the system is obtained from classical PID controller by fuzzy inference generated time varied gains. The nonlinear behavior of the system is characterized from position error, integral of position error and derivative of position error of the controlled system. Therefore, the proposed controller has encountered the angular position control issue without linearizing nonlinear system. As a conclusion, the proposed fuzzy PID controller shows better and satisfactory control performance against external disturbance for nonlinear system compared to classical linear PID controller.

6. References

[1] A. Farmanbordar, N. Zaeri and S. Rahimi, "Stabilizing a Driven Pendulum Using DLQR Control", in *2011 Fifth Asia Modelling Symposium*, 2011, pp. 123-126.
[2] A. Mohammadbagheri and M. Yaghoobi, "A New Approach to Control A Driven Pendulum with PID Method", in *2011 UKSim 13th International Conference on Modelling and Simulation*, 2011, pp. 207-211.
[3] H. Kizmaz, S. Aksoy and A. Muhurcu, “Sliding mode control of suspended pendulum”, Modern Electric Power Systems (MEPS), in *2010 Proceedings of the International Symposium*, 2010, pp. 1-6.
[4] G. Habib, A. Miklos, E. T. Enikov, G. Stepan and G. Rega, “Nonlinear model-based parameter estimation and stability analysis of an aero-pendulum subject to digital delayed control”, *International Journal of Dynamics and Control*, 2015, DOI 10.1007/s40435-015-0203-0.
[5] S. Srinivasulu Raju, T.S. Darshan and B. Nagendra, “Design of Quadratic Dynamic Matrix Control for Driven Pendulum System”, *International Journal of Electronics and Communication Engineering*, Vol. 5, No. 3, pp. 363-370, 2012.
[6] M. Yoon, “Stabilization of a Propeller - Driven Pendulum”, *International Journal of Engineering Research & Technology (IJERT)*, Vol. 5, Iss. 1, pp. 230-233, 2016.
[7] R. Ghasemi, M. R. Rahimi Khoygani, “Designing Intelligent Adaptive Controller for Nonlinear Pendulum Dynamical System”, *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, Vol. 8, No. 11, pp. 2021-2025, 2014.
[8] T. Huba, T. Malatinec, M. Huba, “Propeller-Pendulum for Nonlinear UAVs Control”, *iJOE*, Vol. 9, Issue 1, pp. 42-46, 2013.
[9] E.T. Enikov, V. Polyzoev, J. Gill, “Low-Cost Take-Home Experiment on Classical Control Using Matlab/Simulink Real-Time Windows Target”, in *Proceedings of the 2010 American Society for Engineering Education Zone IV Conference*, 2010, pp. 1-9.
[10] M.R.R. Khoygani, R. Ghasemi, D. Sanaei, “Design Controller for a Class of Nonlinear Pendulum Dynamical System”, *International Journal of Artificial Intelligence*, Vol. 2, No.4, pp. 159-168, 2013.
[11] Y. Gültekin and Y.Taşcıoğlu, “Pendulum Positioning System Actuated by Dual Motorized Propellers”, in *6th International Advanced Technologies Symposium*, 2011, pp. 6-9.
[12] O. Karasakal, E. Yeşil, M. Güzelkaya, I. Eksin, “Implementation of a new self-tuning fuzzy PID controller on PLC”, *Turkish Journal of Electrical Engineering*, Vol. 13, No. 2, pp. 277–286, 2005.
[13] H. B. Kazemian, “Developments of fuzzy PID controllers”, *Expert Systems*, Vol. 22, No. 5, pp. 254-264, 2005.
[14] J.Y. Kim, H.M. Kim, S.K. Kim, J.H. Jeon, H.K. Choi, “Designing an Energy Storage System Fuzzy PID Controller for Microgrid Islanded Operation”, *Energies*, Vol. 4, pp. 1443-1460, 2011.

Note:

Yener Taskin received the BS and MS degrees in 1996 and 1999 from the Department of Mechanical Engineering of Istanbul University and the Ph.D. degree in 2008 from the Department of Mechanical Engineering of Yildiz Technical University, Istanbul. He is currently an Assistant Professor and Head of Mechatronics Division in the Department of Mechanical Engineering of Istanbul University. His research interests include mechatronics, non-linear control theory, modeling and control of vehicle vibrations.

MOBILE VISUAL ACUITY ASSESSMENT APPLICATION: AcuMob

Akhan AKBULUT[1], Muhammed Ali AYDIN[2], Abdül Halim ZAİM[3]

[1]Department of Computer Engineering, Istanbul Kültür University, Istanbul, Turkey
[2]Department of Computer Engineering, Istanbul University, Istanbul, Turkey
[3]Department of Computer Engineering, Istanbul Commerce University, Istanbul, Turkey
a.akbulut@iku.edu.tr, aydinali@istanbul.edu.tr, azaim@ticaret.edu.tr

Abstract: *This paper presents a mobile healthcare (mHealth) system for estimation of visual impairment that provides easiness by specifying the degree of an eye as orthoscopes. Our proposed system called AcuMob which is an Android based mobile application aimed to be used by patients who have myopia. In the crowd society, our proposed app will be implemented faster than the traditional ophthalmologic examination treatments as an alternative. Because AcuMob can be used in everywhere in any time slot, it is offered in the area where the ophthalmologist is not available. The system is developed with using Xamarin framework and voice commands are used to interact with mobile app. Some preferable letters that are suggested by the ophthalmologists were used in the system. The letter categories are specified according to letters' sizes. In the start-up screen, the biggest letter is demonstrated and if the user responds correct answer, the letter's size is being smaller. However, if the user says wrong answer three times consecutively, eyesight ratio is produced by the system to the user referencing to Snellen Chart's information. This article has aimed at making a prediction about the visual impairment's degree. Thanks to AcuMob, people can get idea about their visual acuity without consulting to an eye medical doctor (MD). For the evaluation of systems' reliability, field tests were performed at Bayrampaşa Göz Vakfi Hospital in Istanbul with two ophthalmologist specialists. At the end of trials, the actual diagnosed degrees and the equivalent degree of eyesight ratios according to Snellen Chart's information is compared and the success rates are shown. The system achieved at the 65% of average success rate, which can give users an idea about current condition of their visions.*
Keywords: *m-Health, Mobile application, Visual acuity, Eye test, Eyesight ratio.*

1. Introduction

Visual impairment is a common disease in today's world, and it is mostly seen in people who are aged 65 and older [1]. Although the number of elderly people with visual impairment is much higher, this disease is seen in children with the ratio approximately 26% in this decade [2]. With developing technology, the solutions of the diseases can be found in different platforms and the importance of eHealth has begun to be understood [3].

Ophthalmology is a medical branch that is related to vision problems so that ophthalmologic examination is very common in hospitals in order to diagnose the vision problems [4]. To specifying the degree of vision of people with visual impairment, orthoscopes are used as primary tools.

I this study, a mobile application called AcuMob that grades the eye disorders of patients was presented. Within the proposed system, voice commands were used to interact with the system and Xamarin framework was preferred to implement the application for mobile platforms. To verify the accuracy of the proposed system, patients who were examined by ophthalmologists were asked to use AcuMob after the traditional examination. The prediction rates produced by system and the real diagnostic results coming from ophthalmologists were compared.

The remaining sections of this paper are organized as follows: section two introduces the background and related work, section three explains the approach in detail, section four gives an analysis about test results and section five provides the conclusion and future work.

2. Background and Related Work

Visual impairment is a common disease since the birth of humankind. A huge number of systems were developed in order to find solution to this common condition.

Some of the research that has been developed in this area are summarized in this section.

Tarbert et al. presented a tablet-based application "The Stroke Vision App" for the visual impairment in stroke survivors in order to act as a screening tool. Visual acuity, visual fields and visuospatial neglect can be assessed by this application [5].

Received on: 03.12.2016
Accepted on: 14.03.2017

Lewis et al. developed a simulator with using Unreal Engine 3 game engine. Opticians, visual impairment consultants and group of students tested this virtual environment and test results was promising [6].

In the study of Zhang Xiaomei, the relationship between visual impairment and higher education was investigated. They constructed a network system that includes teaching aids for higher education of people with visual impairment [7].

Another study by Geman et al., a health care self-monitoring system that includes network of sensors transmitting the information was developed. This system warns the users about the obstacles in their way by using aural warnings [8].

Fransis et al. studied the relationship of visually impaired people and their usage of e-commerce web sites. They proposed a framework that will be more suitable for people with visual impairment [9].

In the study of Murphy et al. twenty computer applications that provides touchable sensations were used in the learning phases of math and science classes of visual impaired students [10].

Kii et al. built an accessible optical wireless pedestrian support system that is using a visible light communication with self-illuminated bollards to determine the best distance for danger notification [11].

In the study of Amin et al. a system named Mongol Dip was built to provide audio-based interfaces in the usage of computers [12].

Vlaminck et al. built a drag detection system that provides a 3D atmosphere, which uses multiple sensors to help visual impaired people by increasing their mobility [13].

In the study of Santos et al., a wireless interactive system composed of some modules was designed. This system works with smartphones and embedded systems. Information on the bus stop module is transferred to those who have visual problems for making their transportation easier [14].

Mauro et al. designed a system called DroneNavigator to be used in navigating visually impaired people. They used small insights drones to sense the environment and objects to warn users [15].

In the study of Emiliano et al., visually impaired people can get the information about the environment and routes thanks to their solution Audioguide. The goal of this system is to provide independency to people that have visual impairment during their locomotion [16].

Lisa et al. built an Android application that is used in learning area of geometry for visually impaired students. The system provides an atmosphere to visually impaired students to investigate basic geometric structures showed on a tablet through sound and vibrotactile feedback. A tactile experience is provided by a physical application, which can be a manipulable deformable shape sensed by the tablet [17].

Erin et al. provide a system that contains Wizard-of-Oz navigation interfaces in order to respond to different instruction periods during in situ navigation tasks. They realized an experimental study with nine visual impaired people, and provided them the directions [18].

In the study of Alireza et al. a method was proposed to help people with visual impairment about playing soccer more adequately. The system uses headphone-rendered special audio, a personal computer, and sensors to provide 3D sound representing the objects [19].

3. Methodology

AcuMob was developed to be used for people with visual impairment who can benefit from this application to state their visual acuity anywhere. Voice commands is the only used interaction method for implementing our application. With using to users' voice commands, the system process acuity and states vision problems.

It is planned to use the application from 3 meters to the users. Bluetooth earphones was used to capture voice commands more effectively. Patients are asked to respond correctly to the demonstrated letter within 5 seconds. At the end when patient cannot recognize the screened letter, AcuMob produces a result like "The patient's capacity of see is 70%". As the final step of proving the success of AcuMob, the data coming from the doctors are compared with our application to evaluate the success of prediction.

In today's world, ratio of visual impairment is increasing day by day so that some solutions in order to ease doctor's workload should be found. Our proposed model is designed to be serviced under this goal.

The system was developed using Xamarin framework and Bluetooth earphones was used by the users as supporting device. The vision ratio's information is provided by the Snellen Chart's information.

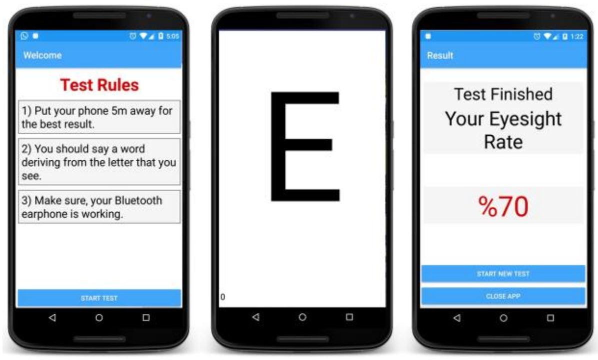


Figure 1. Interfaces of AcuMob

As shown in Figure 1, at the beginning of the application, the system welcomes the user with displaying rules of experiment. Local variables *LetterCategory* and *WrongLetterCounter* were initialized to zero in the start-up. These definitions are used for counting how many times the users respond wrong letters. Then, AcuMob shows the randomly selected first letter and wait 5 seconds to user's perception. After that, the system gives a message that "You can say letter" to user. The decision part that is whether the category of letter is greater than zero or not. If the category of letter is greater than zero, the program outputs a message that is "You see perfectly" that means the user do not have any vision problem. However, if the category of letter is different from zero, another decision part puts into process. This time, if the user does not spell the letter correctly, the system gives a message that is "You did not say correctly" and if this situation repeats three times, the system determines the vision ratio of the user according to values of vision table. If the total number of mislocal letters is less than 3, the system shows another letter with the same size and the user waits 5 seconds and say the letter that is seen from the system just like in the previous steps. Another possibility that, if the user says the letter correctly, the system gives a message about saying the letter successfully, category of letter is planned to be incremented.

Our proposed model is represented in the flowchart below as Figure 2.

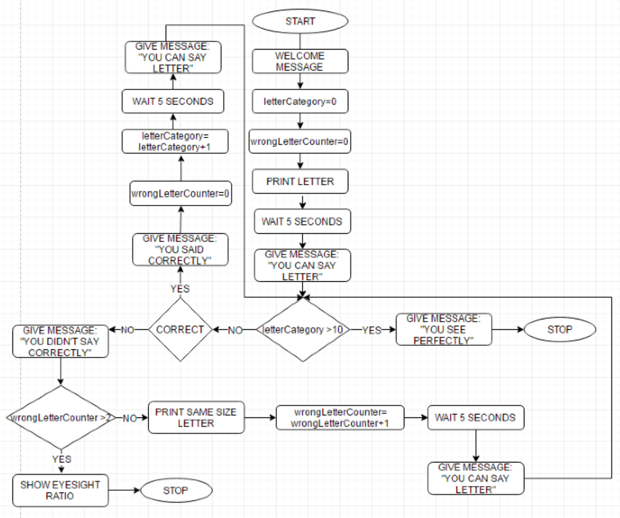


Figure 2. Flowchart of the System

Some letters like A and E are preferred to show for patients to state vision problem. Operator Doctor Şeref Kayabaş from Bayrampaşa Göz Vakfı Hospital from İstanbul recommend these letters to achieve maximum results. Also, as mentioned in other research articles, these letters are frequently used for vision tests [20]. The letters and whose sizes that were used by the program are shown in the Table 1 below.

Table 1. The Letters and Sizes

Letters	Sizes
'E'	160 pt.
'M'	160 pt.
'F'	60, 160 pt.
'U'	105 pt.
'H'	33, 41, 49, 105 pt.
'S'	105 pt.
'N'	80 pt.
'V'	80 pt.
'T'	41, 49, 80 pt.
'Z'	66 pt.
'P'	66 pt.
'D'	66 pt.
'K'	33, 60 pt.
'L'	33, 41, 60 pt.
'A'	49 pt.
'R'	45 pt.
'0'	45 pt.
'Y'	45 pt.

Letters were used in different sizes, because AcuMob provide the users to test the letters' sizes from larger ones to smaller. In the flowchart of the system, letter category was defined to show the sizes of letters. For instance, the letter 'H' is used in sizes that are 33, 41, 49 and 105. For the size 105, letter category is defined as 2. For the size 41, letter category is defined as 8 etc. If the user vocalize the incorrect letter three times consecutively, the system determines the vision ratio. This vision ratio is given according to Snellen Chart that is basic, mostly used and time saving chart in our daily life. It is mostly used from the distance 4 meters to 6 meters [21]. Table 2 shows the Snellen Chart's information and our system's ratios.

Table 2. Snellen Chart and Vision Ratio Information [22]

Snellen Chart Information (Letter's Sizes)	Proposed System's Vision Ratio	Visual Acuity
20/200 ft/m	10%	2.0-2.50 sph.
20/100 ft/m	20%	1.75-2.0 sph.
20/70 ft/m	28% ~ 30%	1.5 sph.
20/50 ft/m	40%	1.0-1.25 sph.
20/40 ft/m	50%	0.5 sph.
20/30 ft/m	60%	0.5 sph.
20/25 ft/m	80%	0.5 sph.
20/20 ft/m	90%	0.0-0.25 sph.

According to Table 2, the biggest letter's size is computed as 20/200 and it has the same size with our system's biggest letter's size that is defined as Letter Category 1. In this situation, the system gives a message to users "Your vision ratio is 10%". For example, the smallest letter's size for system is 33 that is defined as Letter Category 9, so according to Snellen Chart, the smallest letter's size is 20/20 ft/m that is the 1 means the users see perfectly, and the system gives message to users "You see perfectly!".

Proposed system gives a result about user's visual acuity at the end as a percentage. According to Snellen Chart that can be seen in Table 2 above, the proposed system's vision ratio corresponds to user's real visual acuity (eye's degrees) so that users shall get idea about their visual acuity in the neighborhood of.

4. Results and Discussion

We did some field tests to measure the performance of our implementation with visual impaired people in the hospital.

The test procedure was applied like below:

- Primarily, the ophthalmologists examined their patients.
- Then, the ophthalmologists diagnosed the degree of visual impairment of left eyes of patients (1.0, 2.0, etc.).
- Similarly, the ophthalmologists diagnosed the degree of visual impairment of right eyes of patients (1.0, 2.0, etc.).
- Then, the patients who were examined were directed to use ACUMOB application as an alternative approach.
- After that, Bluetooth earphones was used by the patients and the letters were seen in the phone's screen that was five meters away from them.
- Then, the eyesight ratios of left eyes were gathered and the comparison between results coming from

ophthalmologists and the ratios coming from system was realized.

- In the same vein, the eyesight ratios of right eyes were gathered and the comparison between results coming from ophthalmologists and the ratios coming from system was realized.
- Finally, the predicted success rate of the system was evaluated by the doctors.

As detailed in Table 3, the ophthalmologist firstly examined five patients. Their right and left eyes' visual acuities were examined and diagnosed separately. The measurements on the right side of the table show estimates of the same patients by the AcuMob system.

Table 3. Comparison between Patient's Information and Proposed System's Eyesight Ratios

Traditional Examination			**AcuMob Scores**		
Patient 1	1.0	1.0	**Patient 1**	30%	30%
Patient 2	1.50	1.0	**Patient 2**	20%	30%
Patient 3	1.25	1.75	**Patient 3**	50%	40%
Patient 4	1.5	2.0	**Patient 4**	40%	30%
Patient 5	2.50	2.25	**Patient 5**	30%	30%

In Table 4 below, according to the values coming from Table 3, the success rates of right eyes and left eyes' acuities were computed by using proportion technique. For right eyes of these five patients, the total success rates were found as 61.6%, and for left eyes of these patients. The total success rates shoed up as 69.16%. The mean performance of the patients in the first group was determined as 65.38%.

Table 4. Comparison between Patient's Information and Snellen Chart's Estimated Prescription

Traditional Examination			**AcuMob's Readings**			**Success Rates**	
	Right Eye	Left Eye		Right Eye	Left Eye		
Patient 1	1.0	1.0	Patient 1	1.5	1.5	66.6%	66.6%
Patient 2	1.50	1.0	Patient 2	2.0	1.5	75%	66.6%
Patient 3	1.25	1.75	Patient 3	0.5	1.25	40%	71%
Patient 4	1.5	2.0	Patient 4	1.0	1.5	66.6%	75%
Patient 5	2.50	2.25	Patient 5	1.5	1.5	60%	66.6%
					TOTAL	**61.6%**	**69.16%**

In Table 5 below, another group of five patients were examined by the ophthalmologists. Their right and left eyes' visual acuities were examined separately. On the right side of the table, the estimates produced by AcuMon are shown, as in the first experiment.

Table 5. Comparison between Patient's Information and System's Eyesight Ratios

Traditional Examination			**AcuMob Scores**		
Patient 1	0.50	0.75	**Patient 1**	40%	40%
Patient 2	1.5	1.5	**Patient 2**	20%	20%
Patient 3	2.75	2.50	**Patient 3**	30%	30%
Patient 4	1.0	1.25	**Patient 4**	30%	50%
Patient 5	1.75	1.75	**Patient 5**	40%	40%

In Table 6 below, according to the values coming from Table 5, the success rates of right eyes and left eyes' acuities were computed again by using proportion technique. For right eyes of second group, the total success rates were found as 65.2%, and for left eyes, the total success rates were found as 66.2%. The mean performance of the patients in the second group was determined as 65.7%.

Table 6. Comparison between Patient's Information and Snellen Chart's Estimated Prescription

Traditional Examination			**AcuMob's Readings**			**Success Rates**	
	Right Eye	Left Eye		Right Eye	Left Eye		
Patient 1	0.50	0.75	Patient 1	1.0	1.0	50%	75%
Patient 2	1.5	1.5	Patient 2	1.75	1.75	85%	85%
Patient 3	2.75	2.50	Patient 3	1.5	1.5	54%	60%
Patient 4	1.0	1.25	Patient 4	1.5	0.50	66%	40%
Patient 5	1.75	1.75	Patient 5	1.25	1.25	71%	71%
					TOTAL	**65.2%**	**66.2%**

As can be seen from the tables, the success rates of the system are about 60%. This score is truly acceptable to give an idea to the users about their visual acuities from a mobile application. Users were provided with an assessment of eye values in 2 minutes without any medical expert support. It is quite practical to use this app with any Android-based smartphone in any environment. In situations where the environment is silent, it has been experimented that the application works successfully without using headphones.

5. Conclusion

This research has aimed at making a prediction about the visual acuity of users via a mobile application. Thanks to AcuMob system, people can get idea about their visual acuity approximately.

The proposed system was developed on the Xamarin framework and it is designed to be used in Android platform. As part of future work, the necessary arrangements will be made so that AcuMob can also run on the iOS platform.

The proposed system was implemented for the patients who have visual impairment problem that is myopia. In the future, the proposed approach can be developed for the patients who have hypermetropia.

Totally, 18 letters and 9 letter's sizes were used in the interfaces of AcuMob. In the welcome screen of the application, the biggest letter is demonstrated As the user responds correctly, the letters are reduced in size. However, if the user says wrong answer, another letter with same size is seen on the screen. If the user responds three times wrong answer, the system outputs the eyesight ratio to the user. According to Snellen Chart's information, the eyesight ratio can be computed.

For the system reliability, the field tests were performed in Istanbul Bayrampaşa Göz Vakfı Hospital with two ophthalmologist specialists. Firstly, the ophthalmologists examined their patients and the visual acuity of their visual impairment is diagnosed. After that, the patients were asked to try AcuMob program by themselves. Experimental phone was placed five meters away from the patients so that the Bluetooth earphones is used to capture patient's commands.

At the end of trials, the diagnosed degrees and the equivalent degree of eyesight ratios according to Snellen Chart's information is compared and the success rates are shown. AcuMob achieved at the 65% of success rate that is acceptable for an alternative diagnosis, for ten patients.

6. Acknowledgements

We would like to thank to ophthalmologists; Şeref Kayabaş, M.D. and Fatma Seçkin Erdem, M.D. from Bayrampaşa Göz Vakfı Hospital for their contributions. The authors would also like to especially thank Aslı Kapıcıoğlu and Utku Doğanay for their assistance and hard work.

7. References

[1] Laforge,R., Spector,W., Sternberg, J. "The Relationship between vision and hearing impairment to one-year mortality and functional decline". *Journal of Aging and Health* 1992; 4:126-148.

[2] Keeffe, J., & Taylor, H. R. "Visual impairment in children". *The British journal of* ophthalmology 1995; 79:624.

[3] Ferrante, E. "Evolving Telemedicine and e-Health". *Telemedicine and e-Health* 2005; 11(3): 370-383.

[4] Duke-Elder, S. "Textbook of ophthalmology". *British medical journal* 1950; 1:650.

[5] Tarbert, C.M.; Livingstone, I.A.T.; Weir, A.J., "Assessment of visual impairment in stroke survivor's". *Engineering in Medicine and Biology Society* (EMBC), 2014; 2185-2188.

[6] Lewis, J.; Brown, D.; Cranton, W.; Mason, R. "Simulating visual impairments using the Unreal Engine 3 game engine". *Serious Games and Applications for Health* 2011; 1-8.

[7] Xiaomei, Z., Zhen, L.,Gang, Z., Yi,S. "Construction of Teaching Aids Network System for Higher Education of People with Visual Impairment" *Intelligent Networks and Intelligent Systems* (ICINIS) 2012; 344-347.

[8] Geman, O.; Chiuchisan, I. "A health care self-monitoring system for patients with visual impairment using a network of sensors". *E-Health and Bioengineering Conference* (EHB) 2015; 1-4.

[9] Francis, H.; Al-Jumeily, D.; Lund, T.O. "A Framework to Support E-Commerce Development for People with Visual Impairment". *Developments in eSystems Engineering (DeSE)* 2013; 335-341.

[10] Murphy, K.; Darrah, M. "Haptics-Based Apps for Middle School Students with Visual Impairments". *Haptics, IEEE Transactions* 2015; 8:318-326.

[11] Kii, H.; Murata, Y.; Oshiba, S.; Nagai, Y.; Watanabe, H.; Iki, S.; Kitani, Y.; Kuwahara, N.; Morimoto, K. "Accessible Optical Wireless Pedestrian-Support Systems for Individuals with Visual Impairment". *Advanced Applied Informatics (IIAIAAI)* 2014; 776-779.

[12] Amin, M.R.; Paul, B.; Khan, F.A. "Bi-lingual audio assistance supported screen reading software for the people with visual impairments". *Computer and Information Technology (ICCIT)* 2012; 202-208.

[13] Vlaminck, M.; Jovanov, L.; Van Hese, P.; Goossens, B.; Philips, W.; Pizurica, A. "Obstacle detection for pedestrians with a visual impairment based on 3D imaging". *3D Imaging (IC3D)* 2013; 1-7.

[14] Antos, E.A.B. "Design of an interactive system for city bus transport and visually impaired people using wireless communication, smartphone and embedded system". *Microwave and Optoelectronics Conference* 2015;1-5.

[15] Avila,M., Funk, M., Henze, N. DroneNavigator: "Using Drones for Navigating Visually Impaired Persons". *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* 2015; 327-328.

[16] Firmino, E.,Teófilo, M. "Visually impaired navigation assistant for emerging market using tactile floor, feature phone and audio descriptions". *Proceedings of the 4th Annual Symposium on Computing for Development* 2013; 2.

[17] Rühmann,L.,Otero,N., Oakley, I. "A Tangible Tool for Visual Impaired Users to Learn Geometry". *10th International Conference on Tangible, Embedded, and Embodied Interaction* 2016; 577-583.

[18] Brady, E., Sato,D., Ruan,C., Takagi,H., Asakawa,C. "Exploring Interface Design for Independent Navigation by People with Visual Impairments". *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* 2015; 387-388.

[19] Zare,A., McMullen,K., Gardner-McCune, C. "Design of an accessible and portable system for soccer players with visual impairments". *CHI '14 Extended Abstracts on Human Factors in Computing Systems* 2014; 1237-1242.

[20] Kaiser, P. K. "Prospective Evaluation of Visual Acuity Assessment: A Comparison of Snellen Versus ETDRS Charts in Clinical Practice (An AOS Thesis)". *Transactions of the American Ophthalmological Society* 2015; 107: 311–324.

[21] Tewari HK, Kori V, Sony P, Venkatesh P, Garg S. "Snellen chart may be preferable over early treatment diabetic retinopathy study charts for rapid visual acuity assessment". *Indian J Ophthalmol* 2006; 54:214.

[22] Sue, S. "Test distance vision using a Snellen chart". *Community Eye Health* 2007; 20: 52.

Akhan Akbulut received his undergraduate and graduate degrees from Computer Engineering Department of Istanbul Kültür University in 2001 and 2008, respectively. In 2013 he completed his Ph.D. thesis on “Extending Wireless Sensor Networks to Internet using Cloud Computing” in Istanbul University. During his master studies, he worked on medium and large-scale e-government, CRM and ERP projects. He has been working at the Department of Computer Engineering in Istanbul Kültür University as Assistant Professor. His research areas are Software Engineering and Cloud Computing.

Muhammed Ali Aydın received his B.S degree in 2001, M.S degree in Computer Engineering in 2005 from Istanbul Technical University and he holds a doctorate in the same discipline from Istanbul University, received in 2009. Dr. Aydın is currently working as an Assistant Professor in Computer Engineering Department of Istanbul University. His main research interests involve computer networks, cryptography and cyber security.

Abdül Halim Zaim obtain Bachelor's Degree in 1993 from Yıldız Technical University Department of Computer Science Engineering Department of Electrical and Electronics Engineering Faculty. He completed his master's degree in Computer Engineering at Boğaziçi University in 1996 and completed his doctorate in Electrical and Computer Engineering at North Carolina State University (NCSU) in 2001. During his Ph.D., he served as a senior researcher at Alcatel USA for one and a half year. Dr. Zaim received his assistant professor title in 2003, Associate Professor in 2004, Professor in 2010 from Istanbul University, respectively. His main academic research interests are Telecommunication, Computer Networks, Performance Modeling, and Software Development.

MULTILANE TRAFFIC DENSITY ESTIMATION AND TRACKING

Mikail YILAN[1], Mehmet Kemal ÖZDEMİR[1]

[1]Graduate School of Natural and Applied Sciences, İstanbul Şehir University, İstanbul, Turkey
mikailyilan@std.sehir.edu.tr, kemalozdemir@sehir.edu.tr

Abstract: *As the number of vehicles in roads increases, information of traffic density becomes crucial to municipalities for making better decisions about road management and to the environment for reduced carbon emission. Here, the problem of traffic density estimation is addressed when there is continuous influx of vehicle data. First the traffic density is modeled by the clusters of the speed groups that are centered after Kernel Density Estimation technique is implemented for the probability density function of the speed data. Then, empirical cumulative distribution function of data is found by Kolmogorov-Smirnov Test. A peak detection algorithm is used to estimate speed centers of the clusters. Since the estimation model has linear and non-linear components, the estimation of variance values and kernel weights are found by a nonlinear Least Square approach with separation of parameters property. Finally, the tracking of former and latter estimations of a road is calculated by using Scalar Kalman Filtering with scalar state - scalar observation generality level. For all example data sets, the minimum mean square error of kernel weights is found to be less than* 0.002 *while error of mean values is found to be less than* 0.261.

Keywords: *Traffic Density, Kernel Density Estimation, Kolmogorov-Smirnov Test, Nonlinear Least Square, Scalar Kalman Filter.*

1. Introduction

The Estimation and prediction of the traffic density is necessary to prevent citizens from congested traffic. If the decision makers have the knowledge of the current and future traffic reports, municipalities would come up with better solutions against the traffic problem. Also, drivers could have better route options to follow while driving. In doing so, they will spend less gasoline and time, and hence low carbon emission and less air pollution will result in. Although there are many different ways to estimate the traffic density, Kernel Density Estimation (KDE) is one of the best estimation techniques since cars that go on the same road with different speeds on different lanes can be better represented by KDE [1]. On the other hand, parametric and non-parametric approaches form the two types of estimation techniques. Former one has a fixed number of parameters and the latter one has an increasing number of parameters when the training data size becomes larger [2]. Since the estimation of traffic parameters needs to cope with continuously incoming data from the field, KDE, which is non-parametric, is exploited to better describe the problem [3]. Gaussian distribution performs well to represent real-time data, and therefore the sample data is typically modeled as normally distributed. In [4] and [5], which are the initial work of this study, KDE was used to derive the probability density function (PDF) of the received data. As mentioned above, KDE reveals various traffic scenarios accurately and it is helpful for theoretical improvements when it is compared with other methods for the PDF expression. In the aforementioned studies, the cumulative distribution function (CDF) was found by Kolmogorov-Smirnov (KS) test, which is less affected by the existence of outliers when compared with other tests [6].

When traffic density is estimated, its data can be thought of a collection of clusters. The clusters are formed by three parameters: kernel weights that show the corresponding cluster's weight among all available clusters, speed centers, and bandwidths. In our first study [5], kernel weights are estimated by using KDE, KS test, and linear Least Square approaches, while the other two parameters were treated as constants. In our earlier study [4], all of these parameters were taken as non-constants. In the first step of estimations, a peak detection algorithm (PDA) over the smoothed version of the PDF was utilized for the estimation of mean values. Nonlinear LS (NLS) with separation of parameters approach was applied successively to estimate variance values and kernel weights. First, speed center's variance and then its kernel weights were estimated. After these estimations, the next speed center's bandwidth and its kernel weight were estimated, and so on. Linear search method that gives accurate results and Newton-Raphson (N-R) Method that reaches to the solution in a quite shorter time were exploited in the NLS approach [7].

In this paper, an extension to the work in [4] will be presented. For the same road, if new data arrives in addition to the already existing data that is used for the estimation, we adopt the tracking of the estimated parameters instead of

Received on: 30.03.2016
Accepted on: 17.05.2016

using the new large data set and recalculating all the parameters needed to make a fresh estimate. In accomplishing this, first an initial estimate is obtained from the new and small sized data. Then, to get an overall estimate, the old and the new estimates are combined. Hence, a forgetting factor is utilized for the tracking of the old and the new estimates so that more weight is given to the newly arrived data rather than updating them according to their number of samples.

In the next part, in Section II, the model of the system will be verbalized. In Section III, numerical estimations and their tracking will be presented and the system performance will be assessed. In Section IV, in light of these efforts, the deductions will be made. Finally in Section V, a summary of conclusions and future prospective will be given.

2. The Model

The first two subsections of this section includes some equations that are formulated in [4] but we briefly present them here since they are also used for the tracking algorithm adopted in this paper.

2.1. Finding PDF with KDE and Empirical CDF with KS Test

For a given N independent samples, let $x \equiv \{X_1, \dots, X_N\}$ comes from a continuous PDF f, which is defined on X. Gauss KDE can then be defined as follows [8]:

When the mean of each data sample is X_i and the corresponding variance is σ, then the Gauss Kernel PDF is

$$\hat{f}(x;\sigma) = \frac{1}{N}\sum_{i=1}^{N} \varphi(x, X_i; \sigma), \qquad x \in R \tag{1}$$

where

$$\varphi(x, X_i; \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-X_i)^2/(2\sigma)} \tag{2}$$

When the contribution of the Gauss Kernels are different, then (1), can be re-expressed by including kernel weights, α_i as:

$$\hat{f}(x;\sigma,\alpha) = \frac{1}{N}\sum_{i=1}^{N} \alpha_i \varphi(x, X_i; \sigma), \qquad x \in R \tag{3}$$

where

$$0 \le \alpha_i \le 1 \quad \text{and} \quad \sum \alpha_i = 1 \tag{4}$$

Although there are lots of approaches for the representation of a PDF, when KDE is preferred for mapping, representation via KDE would be more suitable for visualization and a better theoretical background [8]. Hence, the smoothed version of data is treated to represent PDF with KDE since one of the essential purposes of KDE is to produce a smooth density surface over a 2-D geographical space [9].

The next step is the expression of empirical CDF and here, KS Test is chosen for the implementation since it is less disturbed by the outliers. KS Test initially detects the difference among the real and the empirical speed values of data set and also how much they are close to each other. Since the distribution of speed values is assumed to be Gaussian, this assumption can be examined in the CDF plots. The methodology is to arrange all datum in the data with an increased order and then to rescale them [10]. As expected, the CDF plot reaches to unity when the last datum is processed.

When F is defined as CDF, $\hat{F}$ would be the empirical CDF and $F(x)$ is counted as equal to $F_0(x)$ as a hypothesis for all $x \in R$ values. Then, KS Test statistics are defined as supremum of the difference between $\hat{F}(x)$ and $F_0(x)$ as described in [10].

2.2. Determination of Speed Centers via PDA and Estimation of Variance and Kernel Weights with Nonlinear LS Method

The clusters in a data set are decided according to speed centers that correspond to different regions in the PDF. Therefore, primarily all mean values should be estimated via PDA. The algorithm is applied to the PDF, and the speed centers are found straightforwardly. First, the derivative of the CDF is calculated that produces the PDF. Later, the resulting values are smoothed to get the PDF for the algorithm implementation. With these, the PDA generates accurate mean values since it examines every datum in the data set and catches the peak values. Corresponding mean values of such peak values are the centers of the clusters. After determination of mean values, their variances and kernel weights are estimated by using separability of parameters property of NLS. In the model, α is linear and σ is nonlinear with respect to the model as seen in (3). The approach is to estimate α via LS in terms of σ and then to estimate σ. In order to achieve this, the following equation (LS error) should be minimized for α [7];

$$J(\sigma,\alpha) = (x - H(\sigma)\alpha)^T (x - H(\sigma)\alpha) \tag{5}$$

Here x corresponds to F values of empirical CDF, and hence F will be used instead of x. The estimation for α is then:

$$\hat{\alpha} = \left(H^T(\alpha)H(\alpha)\right)^{-1} H^T(\alpha)F \tag{6}$$

Then, by replacing $\hat{\alpha}$ into above LS error (5), we get:

$$J(\sigma,\hat{\alpha}) = F^T \left(I - H(\alpha)\left(H^T(\alpha)H(\alpha)\right)^{-1} H^T(\alpha)\right) F \tag{7}$$

Thus, minimization of $J(\sigma,\hat{\alpha})$ is the same as maximization of the following equation over α:

$$\max_{\alpha} \left[F^T H(\alpha)\left(H^T(\alpha)H(\alpha)\right)^{-1} H^T(\alpha)F\right] \tag{8}$$

We apply the above maximization for each speed cluster with the following intermediary variables:

$$H_i = \frac{1}{2}\left(1 + erf\,\frac{x-\mu_i}{\sigma\sqrt{2}}\right) \tag{9}$$

$$F = \sum \alpha_i H_i \tag{10}$$

$$F_i = \left(\alpha_i - \sum_{j=1}^{i-1} \alpha_j\right) H_i \tag{11}$$

where F_i and H_i are for F and H, respectively, for the i^{th} cluster and μ_i is the mean speed value for the same cluster. Then the kernel weights α_is for each cluster can be determined as

$$\alpha_i = (H_i^T H_i)^{-1} H_i^T F_i - \sum_{j=1}^{i-1} \alpha_j \tag{12}$$

Initially, the bandwidth of the mean values is calculated via (8) by using linear search method or Newton-Raphson Method [4]. By then inserting this into (12), the kernel weight of the corresponding speed center is determined. Since the mean values are already found before the application of NLS method, the same procedure for the estimation of the variance and kernel weight is repeated for each speed center, thereby resulting in a successive estimation process.

2.3. Tracking of Traffic Density Estimation

Tracking is very much needed when newly arrived data needs to be processed in addition to the past data. Instead of going back to the initial state of estimation of the parameters by using all the existing and the newly arrived data, the estimation of the parameters is just updated with the arrival of new data. Hence, the final estimates are like reaching a consensus between already estimated parameters and newly estimated ones. Moreover, the importance of the old and new data is not the same for the estimation, because new data has more emphasis on estimation and is seen as more probable to convey the current traffic scenario. Therefore, on the contrary to just reordering estimation results according to their number of samples in data sets, the use of a forgetting factor is necessary to improve the tracking capability in time varying parameter estimation [11]. Forgetting factor can be defined as the concept of forgetting in which older data is gradually scrapped by taking into consideration of more recent information [12]. The main idea behind this concept is to give less weight to older data and more weight to the new one [12].

In this study, Scalar Kalman Filter (KF) has been used for tracking. Its scalar state - scalar observation $(s[n-1], x[n])$ generality level is chosen as an approach. The scalar state and the scalar observation equations are as follows [7]:

$$s[n] = \lambda s[n] + u[n] \quad n \geq 0 \tag{13}$$

$$x[n] = s[n] + w[n] \tag{14}$$

where λ is called the forgetting factor with $0 < \lambda < 1$, $u[n]$ is White Gaussian Noise (WGN) with $u[n] \sim \mathcal{N}(0, \sigma_u^2)$, $w[n] \sim \mathcal{N}(0, \sigma_w^2)$, and $s[-1] \sim \mathcal{N}(\mu_s, \sigma_s^2)$. $w[n]$ differs from WGN only in that its variance is allowed to change in time. Further assumption is the independence of $u[n]$, $w[n]$, and $s[-1]$.

$s[n]$ is estimated based on the data set $\{x[0], x[1], \ldots, x[n]\}$ as n increases, and this process is simply a type of filtering. KF approach calculates the estimator $\hat{s}[n]$ subjected to the estimator for the previous sample $\hat{s}[n-1]$ and thus, it is recursive in nature [7].

With $n \geq 0$, the scalar KF equations (Prediction (Pr), Minimum Prediction MSE (Min Pre MSE), Kalman Gain (KG), Correction (Cr), Minimum MSE (Min MSE), respectively) for tracking are as follows:

Pr: $\hat{s}[n|n-1] = \lambda \hat{s}[n-1|n-1]$ (15)

Min Pr MSE: $M[n|n-1] = \lambda^2 M[n-1|n-1] + \sigma_u^2$ (16)

KG: $K[n] = (M[n|n-1])/(\sigma_w^2 + M[n|n-1])$ (17)

Cr: $\hat{s}[n|n] = \lambda \hat{s}[n|n-1] + K[n](x[n] - \hat{s}[n|n-1])$ (18)

Min MSE: $M[n|n] = (1 - K[n])M[n|n-1]$ (19)

3. Numerical Calculations

In this part, the system will be tested with 3 examples: the first one examines tracking with the change only in speed centers, the second one evaluates the tracking of kernel weights' changes, and finally the last one investigates what happens if the all variables have new different values. To simulate the given scenarios, Data Set 1 is produced and estimated firstly and then used in all three examples. For the first example, Data Set 2 and Data Set 3 are also created. For the second example, tracking of kernel weights is performed by using Data Set 1 and Data Set 4. In the last example, Data Set 1 and Data Set 5 are used for the estimations and tracking.

Before we further proceed, we will state how some parameters in the tracking process are chosen. For example, the forgetting factor λ is calculated as a ratio of number of samples in the first data set and the number of all samples. By using the forgetting factor, the sample numbers are used implicitly in the tracking equation, however, the tracking is simply neither updating the overall estimation according to the number of samples nor giving equal weights to both old and new estimates. As the examples will show, the final estimates are closer to the estimates based on the newly arrived data rather than the estimates from the old data. Assuming that the traffic data is obtained via GPS data, and since the accuracy of GPS data is at least 95% according to GPS Standards [13], error values σ_u^2 and σ_w^2 are assumed to 0.05. For the tracking of mean, instead of $\hat{s}[n-1|n-1]$, Data Set 1's mean estimation and instead of $x[n]$, new data set's mean estimation are used. For the kernel weight's tracking, kernel weight estimations of the aforementioned data sets are taken into calculations. The MMSE (minimum mean square error) of each variable is calculated separately since the system is scalar.

A traffic scenario with 3 speed centers is assumed. The performance of the approach will be assessed for the estimated means, variances, and kernel weights as well as tracking. The assumed scenario for the first data set has the following parameters with the number of samples $N = 10000$:

$$\mu_1 = 50 \quad \mu_2 = 70 \quad \mu_3 = 100$$

$$\sigma_1^2 = 6 \quad \sigma_2^2 = 7 \quad \sigma_3^2 = 5$$

$$\alpha_1 = 0.3 \quad \alpha_2 = 0.5 \quad \alpha_3 = 0.2$$

It is needed to be emphasized that differently from [4], in addition to Gaussian distribution variance values, the contribution of GPS allowable error is added as variance. For the given example, uniformly distributed additional variance values are 2.5, 3.5, and 5, respectively. The estimation of Data Set 1's mean values via a peak detection algorithm are as follows:

$$\hat{\mu}_1 = 50.1489 \quad \hat{\mu}_2 = 69.9231 \quad \hat{\mu}_3 = 100.1197$$

The estimation results are very close to the real values as the MMSE is 0.04240:0424. Also, the MMSE of each speed center's estimation of Data Set 1 are 0.2220, 0.0059, and 0.0143, respectively. The variances and kernel weights are estimated by using two methods as explained in Section II-B. For N-R Method, which reaches accurate results quicker, the estimated values are as follows:

$$\hat{\sigma}_1^2 = 9.5088 \quad \hat{\sigma}_2^2 = 10.2556 \quad \hat{\sigma}_3^2 = 22.8845$$

$$\hat{\alpha}_1 = 0.3038 \quad \hat{\alpha}_2 = 0.4674 \quad \hat{\alpha}_3 = 0.2278$$

When the error values are analyzed, kernel weights and speed centers have less error when compared to variance's values. However, the estimation of variances is an intermediate step before the estimation of the kernel weights. Although variance estimation provides useful information about the traffic density, the speed centers and kernel weights are more critical in assessing multi-lane traffic density. The MMSE of kernel weights is 0.0019 and also the MMSE of each kernel weights of Data Set 1 are 1.4440×10^{-5}, 0.0011, and 0.0008, respectively. For the linear search method, which takes longer time but that generally provides more accurate results, the estimated variances and kernel weights are as follows:

$$\hat{\sigma}_1^2 = 9.5090 \quad \hat{\sigma}_2^2 = 10.2560 \quad \hat{\sigma}_3^2 = 22.8840$$

$$\hat{\alpha}_1 = 0.3002 \quad \hat{\alpha}_2 = 0.4619 \quad \hat{\alpha}_3 = 0.2379$$

with MMSE of 0.0029.

As can be seen from the estimated values, the proposed approach can accurately estimate the targeted parameters as MMSE values are acceptably small for the traffic density estimation.

For the first example, new data sets (Data Set 2 and Data Set 3) are produced with a 5 km/h increase in speed centers while keeping the other parameters unchanged. These data sets have their number of samples as $N = 1000$, and by doing so we will examine the results of tracking by repeating the same procedure. Here, the expectation is that the second tracking would be closer to the speed centers of new data set than the first tracking. For Data Set 2, estimation of mean values and the overall system's corrected speed centers are as follows:

$$\hat{\mu}_1 = 54.9107 \quad \hat{\mu}_2 = 74.8072 \quad \hat{\mu}_3 = 105.2616$$

$$\hat{\mu}_{cor1} = 53.9121 \quad \hat{\mu}_{cor2} = 71.9633 \quad \hat{\mu}_{cor3} = 103.1484$$

The MMSE of mean estimation of Data Set 2 is 0.1561. Again, the estimation is very close to the real values. Also, the MMSE of each speed center's estimation of Data Set 2 are 0.0302, 0.0431, and 0.0828, respectively. As seen from

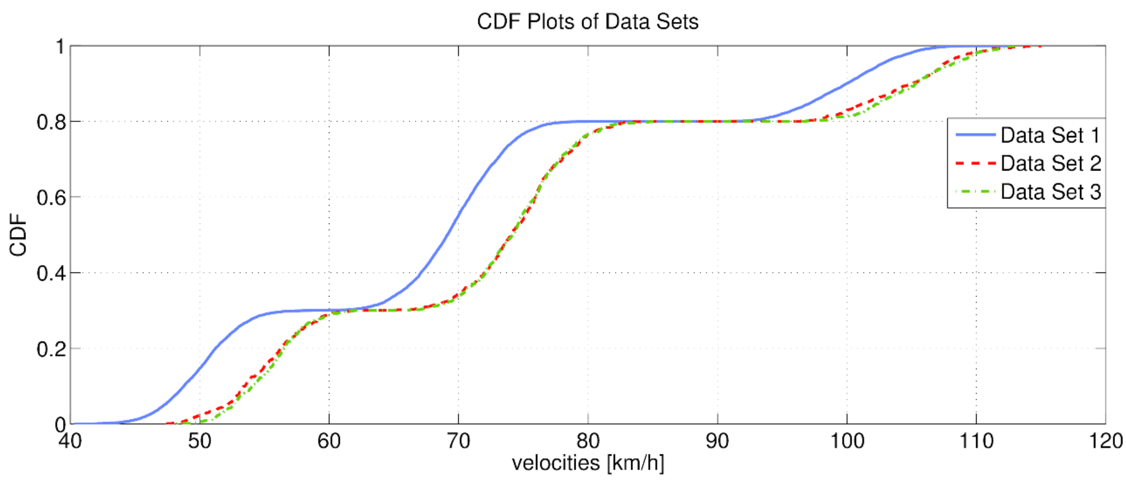


Figure 1. CDF plots of three data sets

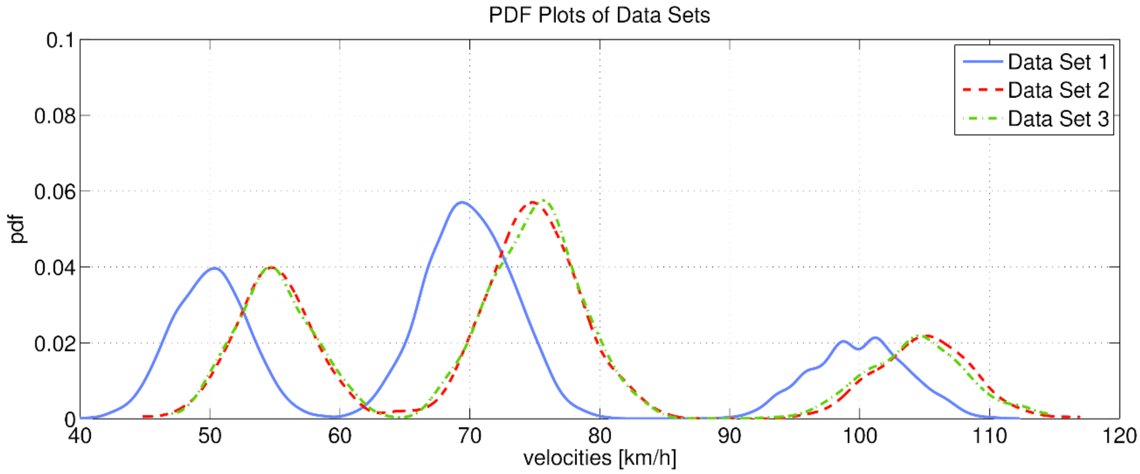


Figure 2. PDF plots of three data sets

results of tracking, corrected mean values are not linearly calculated by simply taking the number of samples in data sets. It is also observed that tracking represents real traffic scenarios better since the new data sets have more effect in the evaluation of the current estimates even though they bear less number of samples. It is obvious that if the initial estimates of the firstly received data are more accurate, then the results are closer to the real values. The estimation of first speed center has more error than the estimation of the second speed center. Thus, the second one has corrected mean values closer to its former estimation (50.1489) than the corrected version of the first mean value's closeness to its former estimation (69.9231). Data Set 3 also has the sameparameter values with Data Set 2. Estimation of its mean values and corrected mean values of the overall system that consists of all three data sets including the results of first tracking are as follows:

$$\hat{\mu}_1 = 54.6585 \quad \hat{\mu}_2 = 75.3391 \quad \hat{\mu}_3 = 105.0514$$

$$\hat{\mu}_{cor1} = 54.2267 \quad \hat{\mu}_{cor2} = 74.7701 \quad \hat{\mu}_{cor3} = 104.6992$$

The MMSE of mean estimation of Data Set 3 is 0.2342. As expected, new corrected values of speed centers are higher than the former ones. The illustration of the change in mean values and their kernel weights and variances of data sets can be observed in Figs. 1 and 2.

In the second example, only kernel weights will change and we will track their values. Data Set 4's kernel weights for $N = 1000$ number of samples are given as follows:

$$\alpha_1 = 0.5 \quad \alpha_2 = 0.4 \quad \alpha_3 = 0.1$$

Since in our case N-R Method's MMSE is less than linear search one, estimation of kernel weights and their tracking are as follows:

$$\hat{\alpha}_1 = 0.5154 \quad \hat{\alpha}_2 = 0.3863 \quad \hat{\alpha}_3 = 0.0983$$

$$\hat{\alpha}_{cor1} = 0.3961 \quad \hat{\alpha}_{cor2} = 0.4027 \quad \hat{\alpha}_{cor3} = 0.1466$$

The MMSE of kernel weights estimation of Data Set 4 is 4.2658×10^{-4}. As seen from corrected kernel weights results, the error of each values of first data set's estimation is correlated with final calculated values. The difference between Data Sets 1 and 4 can be seen in Figs. 3 and 4.

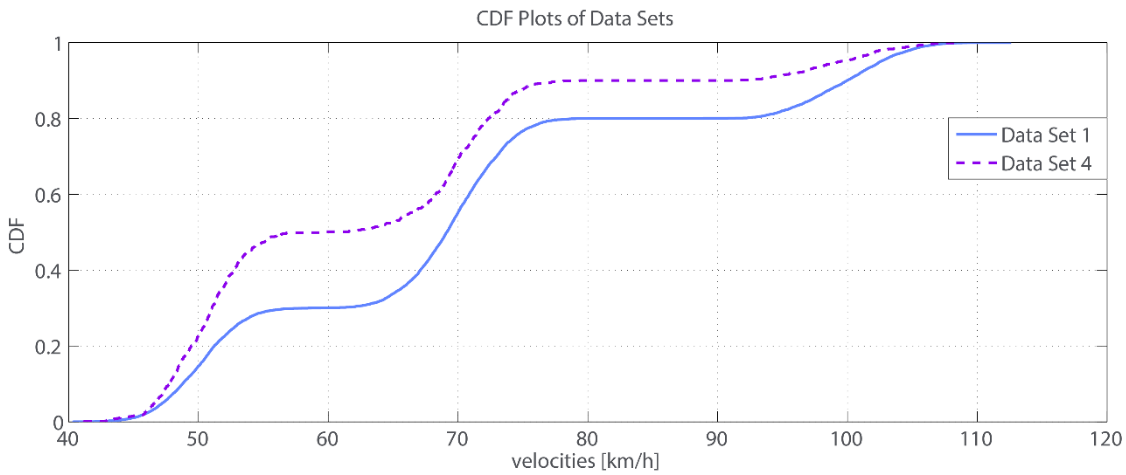


Figure 3. CDF plots of Data Set 1 and Data Set 4

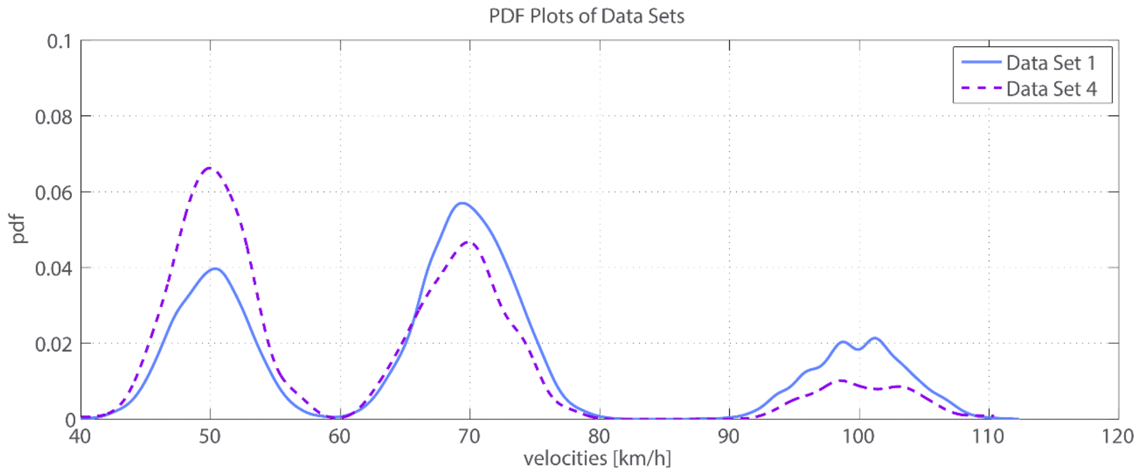


Figure 4. PDF plots of Data Set 1 and Data Set 4

In the last example, all parameters' values of Data Set 1 are changed, and their tracking is calculated for $N = 1000$ sample number. The estimation of mean values via PDA and the estimation of kernel weights and variance values with N-R Method including assumed sample system parameters are given in Table I. Also tracking results of Data Set 5 are shown in the table. The MMSE of mean estimation is 0.2609 and the MMSE of kernel weight estimation is 1.6213×10^{-4}. As seen from all examples, the system has performed well for not only the change of a single parameter but also for the change of all parameters. The PDF and CDF plots of Data Sets 1 and 4 are shown in Figs. 5 and 6.

4. Assessment

The system is tested with three different examples and for all three, it performed well by giving out the desired results. Chosen speed values are relatively middle and high speed levels for driving standards. The model was also examined in low and high speed values, and it performed well for estimations and their tracking. In addition to the work in [4], this work dealt with not only estimations of kernel weights, mean values and variance values but also their tracking. While in estimating variances, N-R Method reaches the results very quickly and its performance is comparable to the linear search method results. For example, for every estimation process of each variance value, linear search method needs more than 100 thousand multiplications to get maximum values in (8), while N-R Method reaches maximum value in less than 10 iteration even though its evaluation needs some heavy computation [4]. Scalar KF with scalar state - scalar observation generality level has achieved preferred outcomes instead of just linear calculations of two estimation results.

Table 1. Example 3 Parameters, Results of Estimation and Tracking

Parameter # \ Name	μ	σ^2	α	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\alpha}$	$\hat{\mu}_{cor}$	$\hat{\alpha}_{cor}$
1	55	5+2.75	0.1	54.7625	6.0102	0.0918	53.7797	0.1838
2	75	6+3.75	0.3	74.9063	12.3964	0.2985	72.0373	0.3523
3	105	7+5.25	0.6	105.4424	19.1120	0.6096	103.3024	0.4331

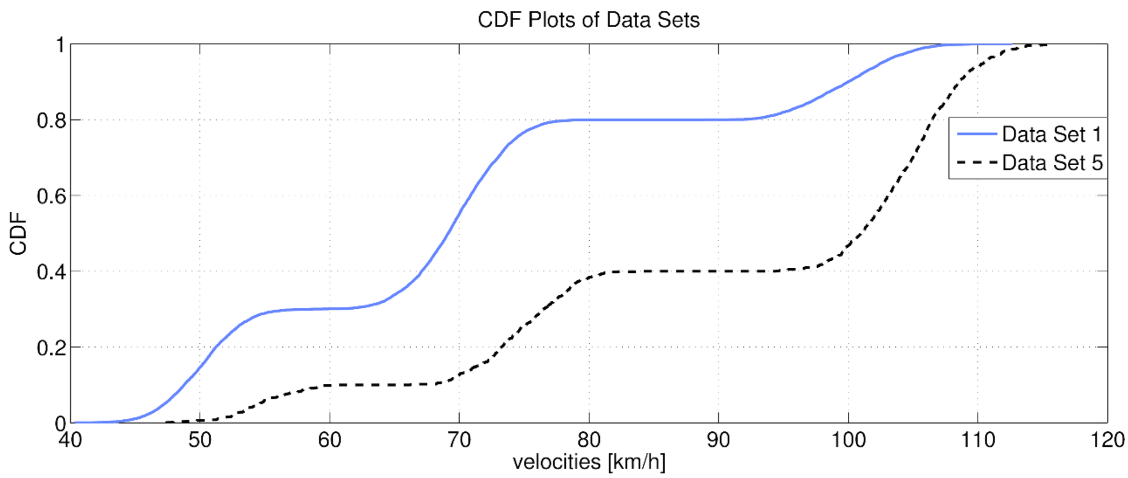


Figure 5. PDF plots of Data Set 1 and Data Set 5

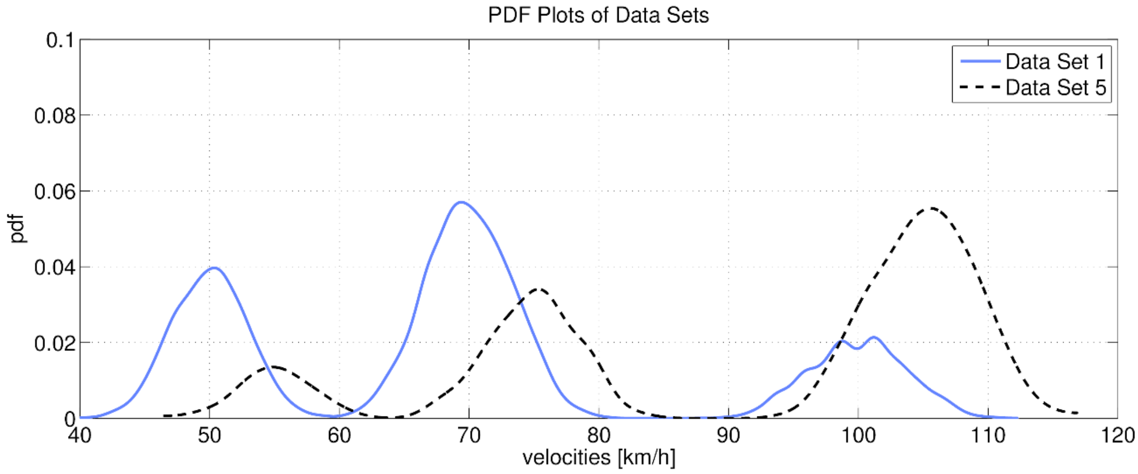


Figure 6. CDF plots of Data Set 1 and Data Set 5

By using both Netwon-Raphson and Linear Search methods, average running time for parameter estimations of old data and new data, and their tracking is approximately 2 minutes and 10 seconds. However, here, estimation of the same parameters are made twice to compare these two methods. As mentioned in Section I, although Newton-Raphson method has computational complexity, it is faster than Linear Search method. If we use only Newton-Rapson method to estimate old data and new data, then perform the tracking, the average running time reduces to approximately 55 seconds. The computer used has a 1.70 GHz CPU, 4.00 GB RAM, and 128.00 GB memory.

The simulation results indicate that the error for mean estimation is less 0.261, while it is less than 0.002 for kernel weights when N-R Method is used. Here the variances are calculated as intermediate variables in estimating the kernel weights. These error rates are considered to represent multilane traffic condition accurately when compared with other studies in the literature such as [14] and [15]. We can use the following mean-percentage error formula that is used in [14] to compare error rates as

$$E_{MPE} = \frac{1}{M}\sum_{i=1}^{M}\left|\frac{\alpha_i \hat{\alpha}_i}{\alpha_i}\right| \tag{14}$$

Here, M is the number of total speed centers and it is equal to 3 in the current study.

The error in [14] is around 13% and the error in [15] is around 10%. Meanwhile, the error rate of mean values in this study reaches a maximum value of 0.11% when (20) is used. If we modify the equation (20) for error rate of kernel weight estimation, its maximum becomes 2.04%. As seen in the Section III, for different data sets and examples (5 data sets for 3 different examples), the error values do not change much and it is sufficiently less than the prior art. Thus, by using the proposed approach, an accurate multi-lane traffic density estimation and its tracking are realized.

5. Conclusion and Future Study

Traffic density estimation and its prediction play a crucial role in managing the traffic on the roads. The overall outcome prevents the drivers from traffic congestion and wasted-time and therefore is very beneficial to both the drivers and the management bodies of municipalities. In this study, multi-lane traffic density estimation has been conducted by estimating the speed centers, bandwidths, and kernel weights of clusters, which represent a group of moving vehicles in a given road and lanes. For this, the PDF of the input data is found by implementing Kernel Density Estimation. Then Kolmogorov-Smirnov Test is used to find empirical CDF. Thereafter, mean values are estimated via a peak detection algorithm and then variance values and kernel weights are estimated successively by using separation of parameters property of nonlinear Least Square Method that is applied with linear search method and Newton-Raphson approaches. As an extension to [4], for the same road, tracking of former and new estimations with less amount of data is determined by using Scalar Kalman Filter with scalar state - scalar observation generality level. The roads' traffic density estimation is then updated with the newly calculated values. Three different sample cases representing a) change in speed centers, b) change in kernel weights, and c) change in all parameters, i.e. mean values, kernel weights, and variance values are analyzed in order to validate the proposed model. It is observed that the proposed estimator and the tracking algorithm perform very well when compared with the state of the art. This current study can further be extended to the prediction of the multi-lane traffic density for a given time interval, say daily or weekly.

6. References

[1] Laxhammar R., Falkman G., and Sviestins E., "Anomaly Detection in Sea Traffic - A Comparison of the Gaussian Mixture Model and the Kernel Density Estimator," in *Information Fusion, 2009, 12th International Conference* on, July 2009, pp. 756–763.
[2] Murphy K. P., *Machine Learning: A Probabilistic Perspective*, MIT press, 2012.

[3] Tabibiazar A. and Basir O., “Kernel-Based Optimization for Traffic Density Estimation in ITS,” in *Vehicular Technology Conference (VTC Fall)*, 2011 IEEE, Sept 2011, pp. 1–5.
[4] Yılan M. and Özdemir M. K., “Traffic Density Estimation via KDE and Nonlinear LS [submitted, pending],” in *Turkish Journal of Electrical Engineering & Computer Sciences*, 2016.
[5] Yılan M. and Özdemir M. K., “A Simple Approach to Traffic Density Estimation by using Kernel Density Estimation,” in *Signal Processing and Communications Applications Conference (SIU), 2015 23th*, May 2015, pp. 1865–1868.
[6] Jäntschi L., Bolboaca S. D. et al., “Distribution Fitting 2. Pearson-Fisher, Kolmogorov-Smirnov, Anderson-Darling, Wilks-Shapiro, Kramer-von-Misses and Jarque-Bera statistics,” *Bulletin UASVM Horticulture*, vol. 66, no. 2, pp. 691–697, 2009.
[7] Kay S. M., *Fundamentals of Statistical Signal Processing: Estimation Theory*, Pearson Education, 2013, vol. 1.
[8] Botev Z. I., Grotowski J. F., Kroese D. P. et al., “Kernel Density Estimation via Diffusion,” *The Annals of Statistics*, vol. 38, no. 5, pp. 2916–2957, 2010.
[9] Xie Z. and Yan J., “Kernel Density Estimation of Traffic Accidents in a Network Space,” *Computers, Environment and Urban Systems*, vol. 32, no. 5, pp. 396–406, 2008.
[10] Djuric P. and Miguez J., “Model Assessment with Kolmogorov-Smirnov Statistics,” in *Acoustics, Speech and Signal Processing, 2009, ICASSP 2009, IEEE International Conference* on, April 2009, pp. 2973–2976.
[11] Song S., Lim J.-S., Baek S., and Sung K.-M., “Gauss Newton Variable Forgetting Factor Recursive Least Squares for Time Varying Parameter Tracking,” *Electronics Letters*, vol. 36, no. 11, pp. 988–990, May 2000.
[12] Vahidi A., Stefanopoulou A., and Peng H., “Recursive Least Squares with Forgetting for Online Estimation of Vehicle Mass and Road Grade: Theory and Experiments,” *Vehicle System Dynamics*, vol. 43, no. 1, pp. 31–55, 2005.
[13] DoD U., “Global Positioning System Standard Positioning Service Performance Standard,” *Assistant Secretary of Defense for Command, Control, Communications, and Intelligence*, 2008.
[14] Munoz L., Sun X., Horowitz R., and Alvarez L., “Traffic Density Estimation with the Cell Transmission Model,” in *American Control Conference, 2003. Proceedings of the 2003*, vol. 5, June 2003, pp. 3750–3755 vol.5.
[15] Sun X., Munoz L., and Horowitz R., “Highway Traffic State Estimation using Improved Mixture Kalman Filters for Effective Ramp Metering Control,” in *Decision and Control, 2003. Proceedings, 42nd IEEE Conference* on, vol. 6, Dec 2003, pp. 6333–6338 Vol. 6.

Note:

Mikail Yılan completed his BSc in electrical-electronics engineering at Boğaziçi University, İstanbul Turkey in '14. He received his MSc from İstanbul Şehir Univesity, Turkey in '16 from electronics and computer engineering. Starting from February '16, he works in industry for IstLink Communications in the area of LTE-A systems. His research interests are traffic density estimation and interference cancellation in HetNets in LTE-A.

Mehmet Kemal Özdemir completed his BSc and MSc in electrical engineering at METU, Ankara Turkey in '96 and in '98, respectively. He received his PhD from Syracuse University, Syracuse, USA in '05 from electrical engineering. Starting from '99, he worked in industry in the area of broadband communication with the focus on CATV and 4G wireless systems. Dr. Ozdemir is currently with Istanbul Sehir University, Istanbul, Turkey as an Asst. Prof at the EE department. His research interests are receiver algorithms for OFDM based systems, traffic density estimation, 5G receiver design, and FM band directional channel modeling.

Novel Conditions for Robust Stability of Bidirectional Associative Memory Neural Networks with Multiple Time Delays

Eylem Yucel

Istanbul University, Department of Computer Engineering,
Istanbul, Turkey, eylem@istanbul.edu.tr

Abstract: *This paper deals with the problem of robust stability of the class of bidirectional associative memory (BAM) neural networks with multiple time delays. Several new sufficient conditions that imply the existence, uniqueness and global robust stability of the equilibrium point for the class of BAM neural networks are obatined by the use of the proper Lyapunov functionals and exploiting the norm properties of the interval matrices. The derived results basically depend on the system parameters of neural network model and they are independent of the time delays. We also give some numerical examples to show the applicability and novelty of the results, and compare the results with the corresponding robust stability results derived in the previous literature.*

Keywords: *Robust Stability, BAM Neural Networks, Lyapunov Theorems, Interval Matrices.*

1. Introduction

In recent years, dynamical neural networks have been extensively studied due to their potential applications in image processing, control theory, pattern recognition, associative memories, optimization problems. In these types of applications, stability properties of the equilibrium point of neural networks are of great importance. In particular, when a neural network is electronically implemented, time delays become important parameters on the stability properties. On the other hand, in hardware implementation of neural networks, the network parameters of the system may change because of the deviations in values of the electronic components. In this case, we need to study the robust stability of neural networks. In the past literature, many different stability results for various neural network models have been reported in [1]-[17]. Bidirectional associative memory (BAM) neural networks have been first introduced in [18]. The stability of the BAM neural networks has been extensively studied in the past years and a great number of various sufficient conditions on the stability of BAM neural networks have been presented in [18]-[35]. However, most of these stability results derived for the BAM neural networks are applicable when neural network model has a single delay. In this paper, we will consider bidirectional associative memory neural networks with multiple time delays. By using some suitable Lyapunov-Krasovskii functionals and properties of intervalized interconnection matrices of the neural system, some new delay-independent sufficient conditions for the existence, uniqueness and global robust asymptotic stability of the equilibrium point for hybrid, BAM neural networks with time delays are derived. Some numerical examples will be presented to show the advantages of our results over to the previous stability results derived in the literature.

2. BAM Neural Networks

Dynamics of a BAM neural network with constant multiple time delays is described by the differential equations of the form :

$$
\begin{aligned}
\dot{u}_i(t) &= -a_i u_i(t) + \sum_{j=1}^{m} w_{ji} g_j\left(z_j(t)\right) + \sum_{j=1}^{m} w_{ji}^{\tau} g_j\left(z_j(t-\tau_{ji})\right) + I_i, \forall i \\
\dot{z}_j(t) &= -b_j z_j(t) + \sum_{i=1}^{n} v_{ij} g_i\left(u_i(t)\right) + \sum_{i=1}^{n} v_{ij}^{\tau} g_i\left(u_i(t-\sigma_{ij})\right) + J_j, \forall j
\end{aligned} \tag{1}
$$

The BAM neural network model (1) consists of two layers. n denotes number of the neurons in the first layer and m denotes the number of neurons in the second layer. $u_i(t)$ is the state of the ith neuron in the first layer and $z_j(t)$ is the state of the jth neuron in the second layer. a_i and b_j denote the neuron charging time constants and passive decay rates, respectively; $w_{ji}, w_{ji}^{\tau}, v_{ij}$ and v_{ij}^{τ} are synaptic connection strengths; g_i and g_j represent the activation functions of the

Received on: 24.05.2016
Accepted on: 31.01.2017

neurons and the propagational signal functions, respectively; and I_i and J_j are the exogenous inputs.

We assume that a_i, b_j, w_{ji}, w_{ji}^{τ}, v_{ij} , v_{ij}^{τ} ,τ_{ji} and σ_{ij} in system (1) are defined at the following intervals:

$$
\begin{aligned}
A_I \ &:= \{A = diag(a_i) : 0 < \underline{A} \le A \le \overline{A}, \\
&\quad i.e., 0 < \underline{a}_i \le a_i \le \overline{a}_i, i = 1,2,\dots,n, \forall A \in A_I\} \\
B_I \ &:= \{B = diag(b_j) : 0 < \underline{B} \le B \le \overline{B}, \\
&\quad i.e., 0 < \underline{b}_j \le b_j \le \overline{b}_j, j = 1,2,\dots,m, \forall B \in B_I\} \\
W_I \ &:= \{W = (w_{ji})_{mxn} : \underline{W} \le W \le \overline{W}, \\
&\quad i.e., \underline{w}_{ji} \le w_{ji} \le \overline{w}_{ji},\ i = 1,2,\dots,n; \\
&\quad j = 1,2,\dots,m, \forall W \in W_I\} \\
V_I \ &:= \{V = (v_{ij})_{nxm} : \underline{V} \le V \le \overline{V}, \\
&\quad i.e., \underline{v}_{ij} \le v_{ij} \le \overline{v}_{ij},\ i = 1,2,\dots,n; \\
&\quad j = 1,2,\dots,m, \forall V \in V_I\} \\
W_I^{\tau} &:= \{W^{\tau} = (w_{ji}^{\tau})_{mxn} : \underline{W}^{\tau} \le W \le \overline{W}^{\tau}, \\
&\quad i.e., \underline{w}_{ji}^{\tau} \le w_{ji}^{\tau} \le \overline{w}_{ji}^{\tau},\ i = 1,2,\dots,n; \\
&\quad j = 1,2,\dots,m, \forall W^{\tau} \in W_I^{\tau}\} \\
V_I^{\tau} &:= \{V^{\tau} = (v_{ij}^{\tau})_{nxm} : \underline{V}^{\tau} \le V \le \overline{V}^{\tau}, \\
&\quad i.e., \underline{v}_{ij}^{\tau} \le v_{ij}^{\tau} \le \overline{v}_{ij}^{\tau},\ i = 1,2,\dots,n; \\
&\quad j = 1,2,\dots,m, \forall V^{\tau} \in V_I^{\tau}\}
\end{aligned}
\tag{2}
$$

The activation functions are assumed to satisfy the following conditions:

(H1) There exist some positive constants ℓ_i, $i = 1,2,\dots,n$ and k_j, $j = 1,2,\dots,m$ such that

$$0 \le \frac{g_i(\overline{x}) - g_i(\overline{y})}{\overline{x} - \overline{y}} \le \ell_i, \qquad 0 \le \frac{g_j(\hat{x}) - g_i(\hat{y})}{\hat{x} - \hat{y}} \le k_j$$

for all $\overline{x}, \overline{y}, \hat{x}, \hat{y} \in R$. This class of functions is denoted by $g \in K$.

(H2) There exist positive constants M_i, $i = 1,2,\dots,n$ and L_j, $j = 1,2,\dots,m$ such that $|g_i(u)| \le M_i$ and $|g_j(z)| \le L_j$ for all $u, z \in R$. This class of functions is denoted by $g \in B$.

3. Preliminaries

Let $v = (v_1, v_2, \dots, v_n)^T \in R^n$ be a column vector and $Q = (q_{ij})_{nxn}$ be a real matrix. The three commonly used vector norms $\|v\|_1, \|v\|_2, \|v\|_\infty$ are defined as :

$$\|v\|_1 = \sum_{i=1}^{n} |v_i|, \|v\|_2 = \sqrt{\sum_{i=1}^{n} v_i^2}, \ \ \|v\|_\infty = \max_{1\le i\le n} |v_i|.$$

The three commonly used matrix norms $\|Q\|_1$, $\|Q\|_2, \|Q\|_\infty$ are defined as follows:

$$\|Q\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |q_{ij}|, \ \ \|Q\|_2 = [\lambda_M(Q^T Q)]^{\frac{1}{2}},$$

$$\|Q\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |q_{ij}|.$$

If $v = (v_1, v_2, \dots, v_n)^T$, then, $|v|$ will denote $v = (|v_1|, |v_2|, \dots, |v_n|)^T$. If $Q = (q_{ij})_{nxn}$, then, $|Q|$ will denote $|Q| = (|q_{ij}|)_{nxn}$, and $\lambda_m(Q)$ and $\lambda_M(Q)$ will denote the minimum and maximum eigenvalues of Q, respectively.

Lemma 1 [36] : Let A be any real matrix defined by $A \in A_I := \{A = (a_{ij}) : \underline{A} \le A \le \overline{A},\ i.e.,\ \underline{a}_{ij} \le a_{ij} \le \overline{a}_{ij}, i,j = 1,2,\dots,n\}$. Define $A^* = \frac{1}{2}(\overline{A} + \underline{A})$ and $A_* = \frac{1}{2}(\overline{A} - \underline{A})$. Let

$$\sigma_1(A) = \sqrt{\|\ |A^{*T} A^*| + 2|A^{*T}|A_* + A_*^T A_*\ \|_2}$$

Then, the following inequality holds

$$\|A\|_2 \le \sigma_1(A)$$

Lemma 2 [37] : Let A be any real matrix defined by $A \in A_I := \{A = (a_{ij}) : \underline{A} \le A \le \overline{A},\ i.e., \underline{a}_{ij} \le a_{ij} \le \overline{a}_{ij}, i,j = 1,2,\dots,n\}$. Define $A^* = \frac{1}{2}(\overline{A} + \underline{A})$ and $A_* = \frac{1}{2}(\overline{A} - \underline{A})$. Let

$$\sigma_2(A) = \|A^*\|_2 + \|A_*\|_2$$

Then, the following inequality holds

$$\|A\|_2 \le \sigma_2(A)$$

Lemma 3 [38] : Let A be any real matrix defined by $A \in A_I := \{A = (a_{ij}) : \underline{A} \le A \le \overline{A},\ i.e., \underline{a}_{ij} \le a_{ij} \le \overline{a}_{ij}, i,j = 1,2,\dots,n\}$. Define $A^* = \frac{1}{2}(\overline{A} + \underline{A})$ and $A_* = \frac{1}{2}(\overline{A} - \underline{A})$. Let

$$\sigma_3(A) = \sqrt{\|\ A^*\|_2^2 + \|A_*\|_2^2 + 2\|A_*^T |A^*|\ \|_2}$$

Then, the following inequality holds

$$\|A\|_2 \le \sigma_3(A)$$

Lemma 4 [39] : Let A be any real matrix defined by $A \in A_I := \{\ A = (a_{ij}) :\ \underline{A} \le A \le \overline{A},\ i.e.,\ \underline{a}_{ij} \le a_{ij} \le \overline{a}_{ij}.\ i,j = 1,2,\dots,n\}$ Define $\hat{A} = (\hat{a}_{ij})_{nxn}$ and $\hat{a}_{ij} = \max\{|\underline{a}_{ij}|, |\overline{a}_{ij}|\}$. Let

$$\sigma_4(A) = \|\hat{A}\|_2$$

Then, the following inequality holds

$$\|A\|_2 \le \sigma_4(A)$$

4. Global Robust Stability Results

In this section, we present some sufficient conditions for the global robust asymptotic stability of the equilibrium point of neural network model (1). First, the equilibrium point of system (1) will be shifted to the origin. By the transformation

$$x_i(.) = u_i(.) - u_i^*, \qquad i = 1, 2, \dots, n,$$
$$y_j(.) = z_j(.) - z_j^*, \qquad j = 1, 2, \dots, m,$$

system (1) can be transformed into a new system of the following form :

$$\dot{x}_i(t) = -a_i x_i(t) + \sum_{j=1}^{m} w_{ji} f_j\left(y_j(t)\right) + \sum_{j=1}^{m} w_{ji}^{\tau} f_j\left(y_j(t-\tau_{ji})\right), \forall i$$
$$\dot{y}_j(t) = -b_j y_j(t) + \sum_{i=1}^{n} v_{ij} f_i\left(x_i(t)\right) + \sum_{i=1}^{n} v_{ij}^{\tau} f_i\left(x_i(t-\sigma_{ij})\right), \forall j \qquad (3)$$

where $x(t) = (x_1(t), x_2(t), \dots, x_n(t))^T$, $y(t) = (y_1(t), y_2(t), \dots, y_n(t)))^T$, $f(x(t)) = (f_1(x_1(t)), f_2(x_2(t)), \dots, f_n(x_n(t)))^T$, $f(y(t)) = (f_1(y_1(t)), f_2(y_2(t)), \dots, f_n(y_n(t)))^T$, $f(x(t-\sigma)) = (f_1(x_1(t-\sigma_1)), f_2(x_2(t-\sigma_2)), \dots, f_n(x_n(t-\sigma_n)))^T$, $f(y(t-\tau)) = (f_1(y_1(t-\tau_1)), f_2(y_2(t-\tau_2)), \dots, f_n(y_n(t-\tau_n)))^T$.

The functions $f_i(x_i), f_j(y_j)$ are of the form :

$$f_i\left(x_i(.)\right) = g_i(x_i(.) + u_i^*) - g_i(u_i^*), i = 1,2, \dots, n,$$
$$f_j\left(y_j(.)\right) = g_j\left(y_j(.) + z_j^*\right) - g_j\left(z_j^*\right), j = 1,2, \dots, m.$$

It can be noted that the functions f_i and f_j satisfy the assumptions on g_i and g_j , i.e., $g_i \in K$ and $g_j \in B$ implies that $f_i \in K$ and $f_j \in B$, respectively. It is also easy to see that $f_i(0) = 0$ and $f_j(0) = 0$, $i = 1,2, \dots, n$.

Note that the equilibrium point of system (1) is globally asymptotically stable, if the origin of system (3) is a globally asymptotically stable. Therefore, the proof of global asymptotic stability of the equilibrium point of system (1) is equivalent to the proof of the global asymptotic stability of the origin of system (3). We now state the following result :

Theorem 1 : Let the assumptions (H1) and (H2) hold. Then, neural system (1) with (2) has a unique equilibrium point which is globally asymptotically robustly stable if there exist positive constants α, γ and β such that

$$\delta_i = m\left(2\underline{a}_i - \alpha - \gamma\right) - \frac{1}{\gamma} n \ell_i^2 \sigma_m^2(V) - \frac{1}{\alpha} n^2 \ell_i^2 \sum_{j=1}^{m} \left(v_{ij}^{\tau^*}\right)^2 > 0, \forall i$$

$$\Omega_j = n\left(2\underline{b}_j - \alpha - \beta\right) - \frac{1}{\beta} m k_j^2 \sigma_m^2(W) - \frac{1}{\alpha} m^2 k_j^2 \sum_{i=1}^{n} \left(w_{ji}^{\tau^*}\right)^2 > 0, \forall j$$

where $W = (w_{ji}), V = (v_{ij})$, $\sigma_m(V) = min\{\sigma_1(V), \sigma_2(V), \sigma_3(V), \sigma_4(V)\}$, $\sigma_m(W) = min\{\sigma_1(W), \sigma_2(W), \sigma_3(W), \sigma_4(W)\}$, $v_{ij}^{\tau^*} = max\{|\underline{v}_{ij}^{\tau}|, |\overline{v}_{ij}^{\tau}|\}$, $w_{ji}^{\tau^*} = max\{|\underline{w}_{ji}^{\tau}|, |\overline{w}_{ji}^{\tau}|\}$

Proof: Define the following positive definite Lyapunov functional :

$$V\left(x(t), y(t)\right) = \sum_{i=1}^{n} m x_i^2(t) + \sum_{j=1}^{m} n y_j^2(t) + \frac{1}{\alpha} \sum_{i=1}^{n} \sum_{j=1}^{m} m^2 (w_{ji}^{\tau})^2 \int_{t-\tau_{ji}}^{t} f_j^2\left(y_j(\eta)\right) d\eta + \frac{1}{\alpha} \sum_{j=1}^{m} \sum_{i=1}^{n} n^2 (v_{ij}^{\tau})^2 \int_{t-\sigma_{ji}}^{t} f_j^2\left(x_i(\xi)\right) d\xi$$

The derivative of $V\left(x(t), y(t)\right)$ along the trajectories of the system is obtained as :

$$\begin{aligned}\dot{V}\left(x(t), y(t)\right) = &-\sum_{i=1}^{n} 2m a_i x_i^2(t) + \sum_{i=1}^{n} \sum_{j=1}^{m} 2m x_i(t) w_{ji} f_j\left(y_j(t)\right) + \sum_{i=1}^{n} \sum_{j=1}^{m} 2m x_i(t) w_{ji}^{\tau} f_j\left(y_j(t-\tau_{ji})\right) \\ &- \sum_{j=1}^{m} 2n b_j y_j^2(t) + \sum_{j=1}^{m} \sum_{i=1}^{n} 2n y_j(t) v_{ij} f_i\left(x_i(t)\right) + \sum_{j=1}^{m} \sum_{i=1}^{n} 2n y_j(t) v_{ij}^{\tau} f_i\left(x_i(t-\sigma_{ij})\right) \\ &+ \frac{1}{\alpha} \sum_{i=1}^{n} \sum_{j=1}^{m} m^2 \left(w_{ji}^{\tau}\right)^2 f_j^2\left(y_j(t)\right) - \frac{1}{\alpha} \sum_{i=1}^{n} \sum_{j=1}^{m} m^2 \left(w_{ji}^{\tau}\right)^2 f_j^2\left(y_j(t-\tau_{ji})\right) \\ &+ \frac{1}{\alpha} \sum_{j=1}^{m} \sum_{i=1}^{n} n^2 \left(v_{ij}^{\tau}\right)^2 f_i^2\left(x_i(t)\right)\end{aligned}$$

$$-\frac{1}{\alpha} \sum_{j=1}^{m} \sum_{i=1}^{n} n^2 \left(v_{ij}^{\tau}\right)^2 f_i^2\left(x_i(t-\sigma_{ij})\right)$$

$$\le -\sum_{i=1}^{n} 2ma_i x_i^2(t) + \sum_{i=1}^{n}\sum_{j=1}^{m} 2mx_i(t)w_{ji}f_j\left(y_j(t)\right) + \sum_{i=1}^{n}\sum_{j=1}^{m} 2mx_i(t)w_{ji}^{\tau}f_j\left(y_j(t-\tau_{ji})\right) - \sum_{j=1}^{m} 2nb_j y_j^2(t) + \sum_{j=1}^{m}\sum_{i=1}^{n} 2ny_j(t)v_{ij}f_i\left(x_i(t)\right) + \sum_{j=1}^{m}\sum_{i=1}^{n} 2ny_j(t)v_{ij}^{\tau}f_i\left(x_i(t-\sigma_{ij})\right) + \frac{1}{\alpha}\sum_{i=1}^{n}\sum_{j=1}^{m} m^2\left(w_{ji}^{\tau}\right)^2 k_j^2 y_j^2(t) - \frac{1}{\alpha}\sum_{i=1}^{n}\sum_{j=1}^{m} m^2\left(w_{ji}^{\tau}\right)^2 f_j^2\left(y_j(t-\tau_{ji})\right) + \frac{1}{\alpha}\sum_{j=1}^{m}\sum_{i=1}^{n} n^2\left(v_{ij}^{\tau}\right)^2 \ell_i^2 x_i^2(t) - \frac{1}{\alpha}\sum_{j=1}^{m}\sum_{i=1}^{n} n^2\left(v_{ij}^{\tau}\right)^2 f_i^2\left(x_i(t-\sigma_{ij})\right) \quad (4)$$

We note the following inequalities :

$$\sum_{i=1}^{n}\sum_{j=1}^{m} 2mx_i(t)w_{ji}f_j\left(y_j(t)\right) = 2mx^T(t)Wf\left(y(t)\right) \le m\beta x^T(t)x(t) + m\frac{1}{\beta}f^T\left(y(t)\right)W^TWf\left(y(t)\right) \le m\beta x^T(t)x(t) + m\frac{1}{\beta}\|W\|_2^2\left\|f\left(y(t)\right)\right\|_2^2 \le m\beta\sum_{i=1}^{n} x_i^2(t) + m\frac{1}{\beta}\|W\|_2^2\sum_{j=1}^{m} k_j^2 y_j^2(t) \quad (5)$$

$$\sum_{j=1}^{m}\sum_{i=1}^{n} 2ny_j(t)v_{ij}f_i\left(x_i(t)\right) = 2ny^T(t)Vf\left(x(t)\right) \le n\gamma y^T(t)y(t) + n\frac{1}{\gamma}f^T\left(x(t)\right)V^TVf\left(x(t)\right) \le n\gamma y^T(t)y(t) + n\frac{1}{\gamma}\|V\|_2^2\left\|f\left(x(t)\right)\right\|_2^2 \le n\gamma\sum_{j=1}^{m} y_j^2(t) + n\frac{1}{\gamma}\|V\|_2^2\sum_{i=1}^{n}\ell_i^2 x_i^2(t) \quad (6)$$

$$\sum_{i=1}^{n}\sum_{j=1}^{m} 2mx_i(t)w_{ji}^{\tau}f_j\left(y_j(t-\tau_{ji})\right) \le \sum_{i=1}^{n}\sum_{j=1}^{m}\alpha x_i^2(t) + \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{1}{\alpha}m^2\left(w_{ji}^{\tau}\right)^2 f_j^2\left(y_j(t-\tau_{ji})\right) = m\alpha\sum_{i=1}^{n} x_i^2(t) + \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{1}{\alpha}m^2\left(w_{ji}^{\tau}\right)^2 f_j^2\left(y_j(t-\tau_{ji})\right) \quad (7)$$

$$\sum_{j=1}^{m}\sum_{i=1}^{n} 2ny_j(t)v_{ij}^{\tau}f_i\left(x_i(t-\sigma_{ij})\right) \le \sum_{j=1}^{m}\sum_{i=1}^{n}\alpha y_j^2(t) + \sum_{j=1}^{m}\sum_{i=1}^{n}\frac{1}{\alpha}n^2\left(v_{ij}^{\tau}\right)^2 f_i^2\left(x_i(t-\sigma_{ij})\right) = n\alpha\sum_{j=1}^{m} y_j^2(t) + \sum_{j=1}^{m}\sum_{i=1}^{n}\frac{1}{\alpha}n^2\left(v_{ij}^{\tau}\right)^2 f_i^2\left(x_i(t-\sigma_{ij})\right) \quad (8)$$

Using (5)-(8) in (4) results in

$$\dot{V}\left(x(t),y(t)\right) \le -\sum_{i=1}^{n} 2ma_i x_i^2(t) + m\beta\sum_{i=1}^{n} x_i^2(t) + m\frac{1}{\beta}\|W\|_2^2\sum_{j=1}^{m} k_j^2 y_j^2(t) - \sum_{j=1}^{m} 2nb_j y_j^2(t) + n\gamma\sum_{j=1}^{m} y_j^2(t) + n\frac{1}{\gamma}\|V\|_2^2\sum_{i=1}^{n}\ell_i^2 x_i^2(t)$$

$$+m\alpha\sum_{i=1}^{n}x_i^2(t)+n\alpha\sum_{j=1}^{m}y_j^2(t)$$
$$+\frac{1}{\alpha}\sum_{i=1}^{n}\sum_{j=1}^{m}m^2\left(w_{ji}^{\tau}\right)^2k_j^2y_j^2(t)$$
$$+\frac{1}{\alpha}\sum_{j=1}^{m}\sum_{i=1}^{n}n^2\left(v_{ij}^{\tau}\right)^2\ell_i^2x_i^2(t)$$

Since $\|W\|_2^2\le\sigma_m^2(W)$, $\|V\|_2^2\le\sigma_m^2(V)$ and $\left(w_{ji}^{\tau}\right)^2\le\left(w_{ji}^{\tau^*}\right)^2$, $\left(v_{ij}^{\tau}\right)^2\le\left(v_{ij}^{\tau^*}\right)^2$

$$\dot V\big(x(t),y(t)\big)\le\sum_{i=1}^{n}\Bigg\{m\big(-2\underline{a}_i+\alpha+\gamma\big)$$
$$+\frac{1}{\gamma}n\ell_i^2\sigma_m^2(V)$$
$$+\frac{1}{\alpha}n^2\ell_i^2\sum_{j=1}^{m}\left(v_{ij}^{\tau^*}\right)^2\Bigg\}x_i^2(t)$$
$$+\sum_{j=1}^{m}\Bigg\{n\big(-2\underline{b}_j+\alpha+\beta\big)$$
$$+\frac{1}{\beta}mk_j^2\sigma_m^2(W)$$
$$+\frac{1}{\alpha}m^2k_j^2\sum_{j=1}^{n}\left(w_{ji}^{\tau^*}\right)^2\Bigg\}y_j^2(t)$$
$$=-\sum_{i=1}^{n}\delta_ix_i^2(t)-\sum_{j=1}^{m}\Omega_jy_j^2(t)$$

Since $\delta_i>0$ for $i=1,2,\dots,n$ and $\Omega_j>0$ for $j=1,2,\dots,m$, it follows that $\dot V\big(x(t),y(t)\big)<0$ for $x(t)\neq 0$ or $y(t)\neq 0$. Hence, by the standard Lyapunov-type theorem in functional differential equations we can conclude that the origin of system (3) is globally asymptotically stable.

Theorem 2 : Let the assumptions (H1) and (H2) hold. Then, neural system (1) with (2) has a unique equilibrium point which is globally asymptotically robustly stable if there exist positive constants α and β such that

$$\varphi_i=m\big(2\underline{a}_i-\alpha\ell_i^2-\gamma\big)-\frac{1}{\gamma}n\ell_i^2\sigma_m^2(V)$$
$$-\frac{1}{\alpha}m^2\sum_{j=1}^{m}\left(w_{ji}^{\tau^*}\right)^2>0,\qquad\forall i$$

$$\vartheta_j=n\big(2\underline{b}_j-\alpha k_j^2-\beta\big)-\frac{1}{\beta}mk_j^2\sigma_m^2(W)$$
$$-\frac{1}{\alpha}n^2\sum_{j=1}^{n}\left(v_{ij}^{\tau^*}\right)^2>0,\qquad\forall j$$

where $W=\left(w_{ji}\right)$, $V=\left(v_{ij}\right)$, $\sigma_m(V)=min\{\sigma_1(V),\sigma_2(V),\sigma_3(V),\sigma_4(V)\}$, $\sigma_m(W)=min\{\sigma_1(W),\sigma_2(W),\sigma_3(W),\sigma_4(W)\}$, $v_{ij}^{\tau^*}=max\{|\underline{v}_{ij}^{\tau}|,|\overline{v}_{ij}^{\tau}|\}$, $w_{ji}^{\tau^*}=max\{|\underline{w}_{ji}^{\tau}|,\ |\overline{w}_{ji}^{\tau}|\}$.

Proof : Define the following positive definite Lyapunov functional :

$$V\big(x(t),y(t)\big)=\sum_{i=1}^{n}mx_i^2(t)+\sum_{j=1}^{m}ny_j^2(t)$$
$$+\alpha\sum_{i=1}^{n}\sum_{j=1}^{m}\int_{t-\tau_{ji}}^{t}f_j^2(y_j(\eta))\,d\eta$$
$$+\alpha\sum_{j=1}^{m}\sum_{i=1}^{n}\int_{t-\sigma_{ij}}^{t}f_i^2(x_i(\xi))\,d\xi$$

The derivative of $V(x(t), y(t))$ along the trajectories of the system is obtained as :

$$\dot V\big(x(t),y(t)\big)=-\sum_{i=1}^{n}2ma_ix_i^2(t)$$
$$+\sum_{i=1}^{n}\sum_{j=1}^{m}2mx_i(t)w_{ji}f_j(y_j(t))$$
$$+\sum_{i=1}^{n}\sum_{j=1}^{m}2mx_i(t)w_{ji}^{\tau}f_j(y_j(t-\tau_{ji}))$$
$$+-\sum_{i=1}^{n}2nb_jy_j^2(t)$$
$$+\sum_{j=1}^{m}\sum_{i=1}^{n}2ny_j(t)v_{ij}f_i(x_i(t))$$
$$+\sum_{j=1}^{m}\sum_{i=1}^{n}2ny_j(t)v_{ij}^{\tau}f_i(x_i(t-\sigma_{ij}))$$
$$+\alpha\sum_{i=1}^{n}\sum_{j=1}^{m}f_j^2(y_j(t))$$
$$-\alpha\sum_{i=1}^{n}\sum_{j=1}^{m}f_j^2(y_j(t-\tau_{ji}))$$
$$+\alpha\sum_{j=1}^{m}\sum_{i=1}^{n}f_i^2(x_i(t))$$
$$-\alpha\sum_{j=1}^{m}\sum_{i=1}^{n}f_i^2(x_i(t-\sigma_{ij}))\qquad(9)$$

We also note that

$$\sum_{i=1}^{n}\sum_{j=1}^{m}2mx_i(t)w_{ji}^{\tau}f_j(y_j(t-\tau_{ji})) \leq \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{1}{\alpha}m^2\left(w_{ji}^{\tau}\right)^2x_i^2(t) + \sum_{i=1}^{n}\sum_{j=1}^{m}\alpha f_j^2(y_j(t-\tau_{ji})) \quad (10)$$

$$\sum_{j=1}^{m}\sum_{i=1}^{n}2ny_j(t)v_{ij}^{\tau}f_i(x_i(t-\sigma_{ij})) \leq \sum_{j=1}^{m}\sum_{i=1}^{n}\frac{1}{\alpha}n^2\left(v_{ij}^{\tau}\right)^2y_j^2(t) + \sum_{j=1}^{m}\sum_{i=1}^{n}\alpha f_i^2(x_i(t-\sigma_{ij})) \quad (11)$$

Using (5), (6), (10) and (11) in (9) leads to :

$$\begin{aligned}\dot{V}\big(x(t),y(t)\big) \leq & -\sum_{i=1}^{n}2ma_ix_i^2(t) + m\beta\sum_{i=1}^{n}x_i^2(t) + m\frac{1}{\beta}\|W\|_2^2\sum_{j=1}^{m}k_j^2y_j^2(t) \\ & -\sum_{j=1}^{m}2nb_jy_j^2(t) + n\gamma\sum_{i=1}^{n}y_j^2(t) + n\frac{1}{\gamma}\|V\|_2^2\sum_{i=1}^{n}l_i^2x_i^2(t) \\ & +\alpha n\sum_{j=1}^{m}k_j^2y_j^2(t) + \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{1}{\alpha}m^2\left(w_{ji}^{\tau}\right)^2x_i^2(t) \\ & +\alpha m\sum_{i=1}^{n}l_i^2x_i^2(t) + \sum_{j=1}^{m}\sum_{i=1}^{n}\frac{1}{\alpha}n^2\left(v_{ij}^{\tau}\right)^2y_j^2(t)\end{aligned}$$

Since $\|W\|_2^2 \leq \sigma_m^2(W), \|V\|_2^2 \leq \sigma_m^2(V)$ and $(w_{ji}^{\tau})^2 \leq \left(w_{ji}^{\tau^*}\right)^2$, $(v_{ij}^{\tau})^2 \leq (v_{ij}^{\tau^*})^2$

$$\begin{aligned}\dot{V}\big(x(t),y(t)\big) \leq & \sum_{i=1}^{n}\left\{m\big(-2\underline{a}_i + \alpha\ell_i^2 + \gamma\big) + \frac{1}{\gamma}n\ell_i^2\sigma_m^2(V) + \frac{1}{\alpha}m^2\sum_{j=1}^{m}\left(w_{ji}^{\tau^*}\right)^2\right\}x_i^2(t) \\ & + \sum_{j=1}^{m}\left\{n\big(-2\underline{b}_j + \alpha k_j^2 + \beta\big) + \frac{1}{\beta}mk_j^2\sigma_m^2(W) + \frac{1}{\alpha}n^2\sum_{i=1}^{n}(v_{ij}^{\tau^*})^2\right\}y_j^2(t) \\ = & -\sum_{i=1}^{n}\varphi_i\, x_i^2(t) - \sum_{j=1}^{m}\vartheta_j\, y_j^2(t)\end{aligned}$$

in which $\dot{V}\big(x(t),y(t)\big) < 0$ for all $x(t) \neq 0$ or $y(t) \neq 0$. Hence, the origin of system (3) is globally asymptotically stable.

The following corollaries are the direct results of Theorems 1 and 2 :

Corollary 1: Let $\underline{a}_m = min\{\underline{a}_i\}$, $\underline{b}_m = min\{\underline{b}_j\}$, $\ell_M = max\{l_i\}$, $k_M = max\{k_j\}$.

$$\phi_i = m\big(2\underline{a}_m - \alpha - \gamma\big) - \frac{1}{\gamma}n\ell_M^2\sigma_m^2(V) - \frac{1}{\alpha}n^2\ell_M^2\sum_{j=1}^{m}(v_{ij}^{\tau^*})^2 > 0, \quad \forall i$$

$$\psi_j = n\big(2\underline{b}_m - \alpha - \beta\big) - \frac{1}{\beta}mk_M^2\sigma_m^2(W) - \frac{1}{\alpha}m^2k_M^2\sum_{i=1}^{n}\left(w_{ji}^{\tau^*}\right)^2 > 0, \quad \forall j$$

where $W = (\omega_{ji})$, $V = (v_{ij})$, $\sigma_m(V) = \min\{\sigma_1(V), \sigma_2(V), \sigma_3(V), \sigma_4(V)\}$, $\sigma_m(W) = \min\{\sigma_1(W), \sigma_2(W), \sigma_3(W), \sigma_4(W)\}$, $v_{ij}^{\tau^*} = max\{|\underline{v}_{ij}^{\tau}|, |\overline{v}_{ij}^{\tau}|\}$ and $w_{ji}^{\tau^*} = max\{|\underline{\omega}_{ji}^{\tau}|, |\overline{\omega}_{ji}^{\tau}|\}$.

Corollary 2: Let $\underline{a}_m = min\{\underline{a}_i\}$, $\underline{b}_m = min\{\underline{b}_j\}$, $\ell_M = max\{\ell_i\}$, $k_M = max\{k_j\}$.

$$\zeta_i = m\big(2\underline{a}_m - \alpha\ell_M^2 - \gamma\big) - \frac{1}{\gamma}n\ell_M^2\sigma_m^2(V) - \frac{1}{\alpha}m^2\sum_{j=1}^{m}\left(w_{ji}^{\tau^*}\right)^2 > 0, \quad \forall i$$

$$\xi_j = n\big(2\underline{b}_m - \alpha k_M^2 - \beta\big) - \frac{1}{\beta}mk_M^2\sigma_m^2(W) - \frac{1}{\alpha}n^2\sum_{i=1}^{n}\left(v_{ij}^{\tau^*}\right)^2 > 0, \quad \forall j$$

where $W=(\omega_{ji})$, $V=(v_{ij})$, $\sigma_m(V)=\min\{\sigma_1(V),\sigma_2(V),\sigma_3(V),\sigma_4(V)\}$, $\sigma_m(W)=\min\{\sigma_1(W),\sigma_2(W),\sigma_3(W),\sigma_4(W)\}$, $v_{ij}^{\tau^*}=max\{|\underline{v}_{ij}^{\tau}|,|\overline{v}_{ij}^{\tau}|\}$ and $w_{ji}^{\tau^*}=max\{|\underline{\omega}_{ji}^{\tau}|,|\overline{\omega}_{ji}^{\tau}|\}$.

Corollary 3: Let $\gamma=\ell_M\sigma_m(V)$, $\beta=k_M\sigma_m(W)$.

$$\phi_i=m(2\underline{a}_m-\alpha)-(m+n)(\ell_M\sigma_m(V))-\frac{1}{\alpha}n^2\ell_M^2\sum_{j=1}^{m}\left(v_{ij}^{\tau^*}\right)^2>0,\ \forall i$$

$$\psi_j=n(2\underline{b}_m-\alpha)-(m+n)(k_M\sigma_m(W))-\frac{1}{\alpha}m^2k_M^2\sum_{i=1}^{n}\left(w_{ji}^{\tau^*}\right)^2>0,\ \forall j$$

where $W=(\omega_{ji})$, $V=(v_{ij})$, $\sigma_m(V)=\min\{\sigma_1(V),\sigma_2(V),\sigma_3(V),\sigma_4(V)\}$, $\sigma_m(W)=\min\{\sigma_1(W),\sigma_2(W),\sigma_3(W),\sigma_4(W)\}$, $v_{ij}^{\tau^*}=max\{|\underline{v}_{ij}^{\tau}|,|\overline{v}_{ij}^{\tau}|\}$ and $w_{ji}^{\tau^*}=max\{|\underline{\omega}_{ji}^{\tau}|,|\overline{\omega}_{ji}^{\tau}|\}$.

Corollary 4: Let $\gamma=\ell_M\sigma_m(V),\beta=k_M\sigma_m(W)$.

$$\zeta_i=m(2\underline{a}_m-\alpha\ell_M^2)-(m+n)(\ell_M\sigma_m(V))-\frac{1}{\alpha}m^2\sum_{j=1}^{m}\left(w_{ji}^{\tau^*}\right)^2>0,\ \forall i$$

$$\xi_j=n(2\underline{b}_m-\alpha k_M^2)-(m+n)(k_M\sigma_m(W))-\frac{1}{\alpha}n^2\sum_{i=1}^{n}\left(v_{ij}^{\tau^*}\right)^2>0,\ \forall j$$

where $W=(\omega_{ji})$, $V=(v_{ij})$, $\sigma_m(V)=\min\{\sigma_1(V),\sigma_2(V),\sigma_3(V),\sigma_4(V)\}$, $\sigma_m(W)=\min\{\sigma_1(W),\sigma_2(W),\sigma_3(W),\sigma_4(W)\}$, $v_{ij}^{\tau^*}=max\{|\underline{v}_{ij}^{\tau}|,|\overline{v}_{ij}^{\tau}|\}$ and $w_{ji}^{\tau^*}=max\{|\underline{\omega}_{ji}^{\tau}|,|\overline{\omega}_{ji}^{\tau}|\}$.

5. Comparisons and Examples

In this section, the results obtained in this paper will be compared with the previous global robust stability results of BAM neural networks derived in the literature. In order to make the comparison precise, first the previous results will be restated :

Corollary 5 [40] **:** Let the activation functions satisfy assumptions (H1) and (H2). Then, neural system (1) with (2) has a unique equilibrium point which is globally asymptotically robustly stable if there exist positive constants α, β and γ such that the network parameters of the system satisfy the following conditions

$$\delta_i=m(2\underline{a}_i-\alpha-\gamma)-\frac{1}{\gamma}n\ell_i^2(\|V^*\|_2^2+\|V_*\|_2^2+2\|V_*^T|V^*|\|_2)-\frac{1}{\alpha}n^2\ell_i^2\sum_{j=1}^{m}\left(v_{ij}^{\tau^*}\right)^2>0,\qquad\forall i>0$$

$$\Omega_j=n(2\underline{b}_j-\alpha-\beta)-\frac{1}{\beta}mk_j^2(\|W^*\|_2^2+\|W_*\|_2^2+2\|W_*^T|W^*|\|_2)-\frac{1}{\alpha}m^2k_j^2\sum_{i=1}^{n}\left(w_{ji}^{\tau^*}\right)^2>0,\qquad\forall j>0$$

where $W=(w_{ji})$, $V=(v_{ij})$, $W^*=\frac{1}{2}(\overline{W}+\underline{W})$, $W_*=\frac{1}{2}(\overline{W}-\underline{W})$, $V^*=\frac{1}{2}(\overline{V}+\underline{V})$, $V_*=\frac{1}{2}(\overline{V}-\underline{V})$, $v_{ij}^{\tau^*}=max\{|\underline{v}_{ij}^{\tau}|,|\overline{v}_{ij}^{\tau}|\}$ and $w_{ji}^{\tau^*}=max\{|\underline{w}_{ji}^{\tau}|,|\overline{w}_{ji}^{\tau}|\}$.

Corollary 6 [40]**:** Let the activation functions satisfy assumptions (H1) and (H2). Then, neural system (1) with (2) has a unique equilibrium point which is globally asymptotically robustly stable if there exist positive constants *α, β* and *γ* such that the network parameters of the system satisfy the following conditions

$$\varphi_i=m(2\underline{a}_i-\alpha\ell_i^2-\gamma)-\frac{1}{\gamma}n\ell_i^2(\|V^*\|_2^2+\|V_*\|_2^2+2\|V_*^T|V^*|\|_2)-\frac{1}{\alpha}m^2\sum_{j=1}^{m}\left(w_{ji}^{\tau^*}\right)^2>0,\forall i$$

$$\vartheta_j=n(2\underline{b}_j-\alpha k_j^2-\beta)-\frac{1}{\beta}mk_j^2(\|W^*\|_2^2+\|W_*\|_2^2+2\|W_*^T|W^*|\|_2)-\frac{1}{\alpha}n^2\sum_{i=1}^{n}\left(v_{ij}^{\tau^*}\right)^2>0,\qquad\forall j$$

where $W=(w_{ji})$, $V=(v_{ij})$, $W^*=\frac{1}{2}(\overline{W}+\underline{W})$, $W_*=\frac{1}{2}(\overline{W}-\underline{W})$, $V^*=\frac{1}{2}(\overline{V}+\underline{V})$, $V_*=\frac{1}{2}(\overline{V}-\underline{V})$, $v_{ij}^{\tau^*}=max\{|\underline{v}_{ij}^{\tau}|,|\overline{v}_{ij}^{\tau}|\}$ and $w_{ji}^{\tau^*}=max\{|\underline{w}_{ji}^{\tau}|,|\overline{w}_{ji}^{\tau}|\}$.

We can write the following results for Corollary 5 and Corollary 6 :

Corollary 7: Let $\ell_M=max\{\ell_i\}$, $k_M=max\{k_j\}$, $\gamma=\ell_M\sqrt{\|V^*\|_2^2+\|V_*\|_2^2+2\|V_*^T|V^*|\|_2}$, $\beta=k_M\sqrt{\|W^*\|_2^2+\|W_*\|_2^2+2\|W_*^T|W^*|\|_2}$.

$$\delta_i=m(2\underline{a}_i-\alpha)-(m+n)(\ell_M\sqrt{\|V^*\|_2^2+\|V_*\|_2^2+2\|V_*^T|V^*|\|_2})-\frac{1}{\alpha}n^2\ell_M^2\sum_{j=1}^{m}\left(v_{ij}^{\tau^*}\right)^2>0,\qquad\forall i$$

$$\Omega_j = n(2\underline{b}_j - \alpha) -(m+n)(k_M\sqrt{\|W^*\|_2^2 + \|W_*\|_2^2 + 2\|W_*^T|W^*|\|_2}) - \frac{1}{\alpha}m^2k_M^2\sum_{i=1}^{n}\left(w_{ji}^{\tau^*}\right)^2 > 0, \quad \forall j$$

Corollary 8: Let $\ell_M = max\{\ell_i\}$, $k_M = max\{k_j\}$, $\gamma = \ell_M\sqrt{\|V^*\|_2^2 + \|V_*\|_2^2 + 2\|V_*^T|V^*|\|_2}$, $\beta = k_M\sqrt{\|W^*\|_2^2 + \|W_*\|_2^2 + 2\|W_*^T|W^*|\|_2}$.

$$\varphi_i = m(2\underline{a}_i - \alpha\ell_M^2) -(m+n)\left(\ell_M\sqrt{\|V^*\|_2^2 + \|V_*\|_2^2 + 2\|V_*^T|V^*|\|_2}\right) - \frac{1}{\alpha}m^2\sum_{j=1}^{m}\left(w_{ji}^{\tau^*}\right)^2 > 0, \forall i$$

$$\vartheta_j = n(2\underline{b}_j - \alpha k_M^2) -(m+n)(k_M\sqrt{\|W^*\|_2^2 + \|W_*\|_2^2 + 2\|W_*^T|W^*|\|_2}) - \frac{1}{\alpha}n^2\sum_{i=1}^{n}\left(v_{ij}^{\tau^*}\right)^2 > 0, \forall j$$

Example 1: Assume that the network parameters of neural system (1) are given as follows :

$$\underline{W} = \underline{V} = \begin{bmatrix} 0 & 2a & 2a & 2a \\ -2a & -2a & 2a & 2a \\ 2a & -2a & 2a & -2a \\ -2a & 2a & 2a & -2a \end{bmatrix},$$

$$\overline{W} = \overline{V} = \begin{bmatrix} 0 & 2a & 2a & 2a \\ -2a & -2a & 2a & 2a \\ 2a & -2a & 2a & -2a \\ -2a & 2a & 2a & 0 \end{bmatrix},$$

$$\underline{W}^T = \underline{V}^T = \begin{bmatrix} -2a & -2a & -2a & -2a \\ -2a & -2a & -2a & -2a \\ -2a & -2a & -2a & -2a \\ -2a & -2a & -2a & -2a \end{bmatrix}$$

$$\overline{W}^T = \overline{V}^T = \begin{bmatrix} 2a & 2a & 2a & 2a \\ 2a & 2a & 2a & 2a \\ 2a & 2a & 2a & 2a \\ 2a & 2a & 2a & 2a \end{bmatrix}$$

$\underline{A} = A = \overline{A} = \underline{B} = B = \overline{B} = I$,
$\ell_1 = \ell_2 = \ell_3 = \ell_4 = k_1 = k_2 = k_3 = k_4 = 1$,
Where $a > 0$ is real number. The matrices W^*, W_*, V^*, V_*, $W_*^T|W^*|$, $V_*^T|V^*|$, $\hat{W}$ and $\hat{V}$ are obtained as follows

$$W^* = V^* = \begin{bmatrix} 0 & 2a & 2a & 2a \\ -2a & -2a & 2a & 2a \\ 2a & -2a & 2a & -2a \\ -2a & 2a & 2a & -a \end{bmatrix},$$

$$W_* = V_* = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a \end{bmatrix},$$

$$W_*^T|W^*| = V_*^T|V^*| = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2a^2 & 2a^2 & 2a^2 & a^2 \end{bmatrix},$$

$$\hat{W} = \hat{V} = \begin{bmatrix} 0 & 2a & 2a & 2a \\ 2a & 2a & 2a & 2a \\ 2a & 2a & 2a & 2a \\ 2a & 2a & 2a & 2a \end{bmatrix}$$

We calculate

$$\sigma_1(V) = \sqrt{\||V^{*T}V^*| + 2|V^{*T}|V_* + V_*^TV_*\|_2} = 5{,}0364a$$
$$\sigma_2(V) = \|V^*\|_2 + \|V_*\|_2 = 5{,}8399a$$
$$\sigma_3(V) = \sqrt{\|V^*\|_2^2 + \|V_*\|_2^2 + 2\|V_*^T|V^*|\|_2} = 5{,}6245a$$
$$\sigma_4(V) = \left\|\hat{V}\right\|_2 = 7{,}1231a$$

$\sigma_1(W) = \sigma_1(V)$, $\sigma_2(W) = \sigma_2(V)$, $\sigma_3(W) = \sigma_3(V)$, $\sigma_4(W) = \sigma_4(V)$. Hence

$$\sigma_m(V) = \min\{\sigma_1(V), \sigma_2(V), \sigma_3(V), \sigma_4(V)\} = 5{,}0364a$$
$$\sigma_m(W) = \min\{\sigma_1(W), \sigma_2(W), \sigma_3(W), \sigma_4(W)\} = 5{,}0364a$$

For the network parameters of this example, the conditions of Corollary 3 and Corollary 4 are obtained as follows:

$$\phi_1 = \phi_2 = \phi_3 = \phi_4 = \psi_1 = \psi_2 = \psi_3 = \psi_4 = \zeta_1 = \zeta_2 = \zeta_3 = \zeta_4 = \xi_1 = \xi_2 = \xi_3 = \xi_4 = 8 - 4\alpha - 8(5{,}0364a) - \frac{256a^2}{\alpha}$$

Let $\alpha = 8a$. Hence, if $a < \frac{8}{104{,}2912}$ holds, then the conditions of Corollaries 3 and 4 are satisfied.

We will now check the results of Corollary 7 and Corollary 8 for the same network parameters. The conditions of Corollary 7 and Corollary 8 are obtained as follows:

$$\delta_1 = \delta_2 = \delta_3 = \delta_4 = \Omega_1 = \Omega_2 = \Omega_3 = \Omega_4 = \varphi_1 = \varphi_2 = \varphi_3 = \varphi_4 = \vartheta_1 = \vartheta_2 = \vartheta_3 = \vartheta_4 = 8 - 4\alpha - 8(5{,}6245a) - \frac{256a^2}{\alpha}$$

Let $\alpha = 8a$. Hence, if $a < \frac{8}{108{,}996}$ holds, then the conditions of Corollaries 7 and 8 are satisfied.

Remark: For the parameters in this example, our results require that $a < \frac{8}{104{,}2912}$. However, the results of Corollaries 7 and 8 require that $a < \frac{8}{108{,}996}$. Therefore, for $\frac{8}{108{,}996} \le a < \frac{8}{104{,}2912}$, our conditions obtained in Corollary 3 and Corollary 4 are satisfied but the results of Corollary 7 and Corollary 8 do not hold.

6. Conclusions

In this paper, by using the Lyapunov stability theorems and the norm properties of the interconnection matrices of the neural system, some novel sufficient conditions for the existence, uniqueness and the global robust asymptotic stability of the equilibrium point have been obtained for the class of bidirectional associative memory (BAM) neural networks with multiple time delays. We have also compared our results with the most recent corresponding stability results, implying that our results establish a new set of global robust asymptotic stability criteria for BAM neural networks with multiple time delays.

7. References

[1] C-D. Zheng, H. Zhang and Z. Wang, “Novel Exponential Stability Criteria of High-Order Neural Networks With Time-Varying Delays”, *IEEE Transactions On Systems Man And Cybernetics Part B:Cybernetics* , vol. 41, no. 2, pp. 486-496, 2011

[2] Q. Song and Z. Wang, “Neural networks with discrete and distributed time-varying delays: A general stability analysis”, *Chaos, Solitons and Fractals*, vol. 37, no. 5, pp. 1538-1547, 2008.

[3] Y. He, G.P. Liu, D. Rees and M. Wu, “Stability analysis for neural networks with time-varying interval delay”, *IEEE Transactions On Neural Networks*, vol. 18, no. 6, pp. 1850-1854, 2007.

[4] X. Meng, M. Tian and S. Hu, “Stability analysis of stochastic recurrent neural networks with unbounded time-varying delays”, *Neurocomputing*, vol. 74, no. 6, pp. 949-953, 2011.

[5] H. Zhang, Z. Wang, and D. Liu, “Global asymptotic stability of recurrent neural networks with multiple time varying delays”, *IEEE Transactions on Neural Networks*, vol. 19, no. 5, pp. 855873, 2008.

[6] R. Yang, Z. Zhang and P. Shi, “Exponential Stability on Stochastic Neural Networks With Discrete Interval and Distributed Delays”, *IEEE Transactions on Neural Networks*, vol. 21, no. 1, pp. 169-175, 2010.

[7] R.-S. Gau, C.-H. Lien and J.-G. Hsieh, “Novel Stability Conditions For Interval Delayed Neural Networks With Multiple Time-Varying Delays”, *International Journal Of Innovative Computing Information And Control*, vol. 7, no. 1, pp. 433-444, 2011.

[8] Z. Liu, H. Zhang and Q. Zhang, “Novel Stability Analysis for Recurrent Neural Networks with Multiple Delays via Line Integral-Type L-K Functional”, *IEEE Transactions on Neural Networks*, vol. 21, no. 11, pp. 1710-1718, 2010.

[9] Z. Zuo, C. Yang and Y. Wang, “A New Method for Stability Analysis of Recurrent Neural Networks With Interval Time-Varying Delay”, *IEEE Transactions on Neural Networks*, vol. 21, no. 2, pp. 339-344, 2010.

[10] Z.-G. Wu, Ju H. Park, H. Su and J. Chu, “New results on exponential passivity of neural networks with time-varying delays”, *Nonlinear Analysis: Real World Applications*, vol. 13, no. 4, pp. 1593-1599, 2012.

[11] D.H. Ji, J.H. Koo, S.C. Won, S.M. Lee and Ju H. Park, “Passivity-based control for Hopfield neural networks using convex representation”, *Applied Mathematics and Computation*, vol. 217, no. 13, pp. 6168-6175, 2011.

[12] P. Balasubramaniam and S. Lakshmanan, “Delay-range dependent stability criteria for neural networks with Markovian jumping parameters”, *Nonlinear Analysis: Hybrid Systems*, vol. 3, no. 4, pp. 749-756, 2009.

[13] H. Zhang, Z. Wang and D. Liu, “Global Asymptotic Stability and Robust Stability of a Class of Cohen-Grossberg Neural Networks With Mixed Delays”, *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 3, pp. 616-629, 2009.

[14] P. Balasubramaniam and M.S. Ali, “Robust stability of uncertain fuzzy cellular neural networks with time-varying delays and reaction diffusion terms”, *Neurocomputing*, vol. 74, no. 1-3; pp. 439-446, 2010.

[15] W.-H. Chen and W.X. Zheng, “Robust Stability Analysis for Stochastic Neural Networks With Time-Varying Delay”, *IEEE Transactions on Neural Networks*, vol. 21, no. 3, pp. 508-514, 2010.

[16] Y. Zhao, L. Zhang, S. Shen and H. Gao, “Robust Stability Criterion for Discrete-Time Uncertain Markovian Jumping Neural Networks with Defective Statistics of Modes Transitions”, *IEEE Transactions on Neural Networks*, vol. 22, no. 1, pp. 164-170, 2011.

[17] P. Balasubramaniam, S. Lakshmanan and R. Rakkiyappan, “Delay-interval dependent robust stability criteria for stochastic neural networks with linear fractional uncertainties”, *Neurocomputing*, vol. 72, no. 16-18, pp. 3675-3682, 2009.

[18] B. Kosko, “Adaptive bi-directional associative memories”, *Appl. Opt.*, vol. 26, pp. 4947-4960, 1987.

[19] G. Mathai, B.R. Upadhyaya, “Performance analysis and application of the bidirectional associative memoryto industrial spectral signatures”, *Proc. IJCNN*, vol. 89, no. 1, pp. 33-37, 1989.

[20] S. Arik, “Global Asymptotic Stability Analysis of Bidirectional Associative Memory Neural Networks with Time Delays”, *IEEE Transactions on Neural Networks*, vol. 16, no. 3, pp. 580586, 2005.

[21] JD. Cao, JL. Liang and J. Lam, “Exponential stability of high-order bidirectional associative memory neural networks with time delays”, *Physica D-Nonlinear Phenomena*, vol. 199, no. 3-4, pp. 425-436, 2004.

[22] Z.-T. Huang, . X.-S Luo and Q.-G. Yang, “Global asymptotic stability analysis of bidirectional associative memory neural networks with distributed delays and impulse”, *Chaos, Solitons and Fractals*, vol. 34, no. 3, pp. 878-885, 2007.

[23] X. Lou, B. Cui and W. Wu, “On global exponential stability and existence of periodic solutions for BAM neural networks with distributed delays and reactiondiffusion terms”, *Chaos, Solitons and Fractals*, vol. 36, no. 4, pp. 1044-1054, 2008.

[24] J.H. Park, S.M. Lee and O.M. Kwon, “On exponential stability of bidirectional associative memory neural networks with time-varying delays”, *Chaos, Solitons and Fractals*, vol. 39, pp. 10831091, 2009.

[25] Y. Wang, “Global exponential stability analysis of bidirectional associative memory neural networks with time-varying delays”, *Nonlinear Analysis: Real World Applications*, vol. 10, pp. 1527-1539, 2009.

[26] Y. Yuan and X. Li, “New results for global robust asymptotic stability of BAM neural networks with time-varying delays”, *Neurocomputing*, vol. 74, no. 1-3, pp. 337-342, 2010.

[27] B. Chen, L. Yu and W.-A. Zhang , “Exponential convergence rate estimation for neutral BAM neural networks with mixed time-delays”, *Neural Computing and Applications*, vol. 20, no. 3, pp. 451-460, 2011.

[28] P. Balasubramaniam and C. Vidhya, “Global asymptotic stability of stochastic BAM neural networks with distributed delays and reaction-diffusion terms”, *Journal of Computational and Applied Mathematics*, vol. 234, no. 12, pp. 3458-3466, 2010.

[29] Ju H. Park, C.H. Park, O.M. Kwon and S.M. Lee, "A new stability criterion for bidirectional associative memory neural networks of neutral-type", *Applied Mathematics and Computation*, vol. 199, no. 2, pp. 716-722, 2008.
[30] Ju H. Park, "Robust stability of bidirectional associative memory neural networks with time delays", *Physics Letters A*, vol. 349, no. 6, pp. 494-499, 2006.
[31] X. F. Liao and K. Wong, "Global exponential stability of hybrid bidirectional associative memory neural networks with discrete delays", *Physical Review E*, vol. 67, no. 4, (0402901), 2003.
[32] X. F. Liao and K. Wong, "Robust stability of interval bidirectional associative memory neural network with time delays", *IEEE Trans. Systems, Man and Cybernetics-Part C*, vol. 34, pp. 1142-1154, 2004.
[33] S. Senan and S. Arik, "New results for global robust stability of bidirectional associative memory neural networks with multiple time delays", *Chaos, Solitons and Fractals*, vol. 41, no. 4, pp. 2106-2114, 2009.
[34] S. Senan and S. Arik, "Global robust stability of bidirectional associative memory neural networks with multiple time delays", *IEEE Trans. Systems, Man and Cybernetics-Part B*, vol. 37, no. 5, pp. 1375-1381, 2007.
[35] N.Ozcan and S.Arik, "A new sufficient condition for global robust stability of bidirectional associative memory neural networks with multiple time delays", *Nonlinear Analysis:Real World Applications*, vol.10, pp. 3312-3320, 2009.
[36] O. Faydasicok and S. Arik, S. " A new upper bound for the norm of interval matrices with application to robust stability analysis of delayed neural networks", *Neural Networks*, vol. 44, pp. 6471, 2013.
[37] A. Chen, J. Cao and L. Huang, "Global robust stability of interval cellular neural networks with time-varying delays", *Chaos, Solitons and Fractals*, vol. 23, no :3, pp. 787799, 2005.
[38] T. Ensari and S. Arik, "New results for robust stability of dynamical neural networks with discrete time delays", *Expert Systems with Applications*, vol. 37, no : 8, pp. 59255930, 2010.
[39] V. Singh, "Global robust stability of delayed neural networks: estimating upper limit of norm of delayed connection weight matrix", *Chaos, Solitons and Fractals*, vol. 32, no :1, pp. 259263, 2007.
[40] S. Senan, S. Arik and D. Liu, " New robust stability results for bidirectional associative memory neural networks with multiple time delays", *Applied Mathematics and Computation*, vol. 218, no : 23, pp. 11472-11482, 2013.

Eylem Yücel received the B.Sc., M.Sc. and Ph.D. degrees from Istanbul University, Istanbul, Turkey, in 2001, 2005 and 2010 respectively. She is working as an Assistant Professor at the Department of Computer Engineering, Istanbul University since 2010. Her research interests are neural networks and nonlinear systems.

SOLVING SUDOKU PUZZLE with NUMBERS RECOGNIZED by USING ARTIFICIAL NEURAL NETWORKS

Selcuk SEVGEN, Emel ARSLAN, Ruya SAMLI

Department of Computer Engineering, Istanbul University, Istanbul, Turkey
{sevgens, earslan, ruyasamli}@istanbul.edu.tr

Abstract: *This paper proposed a method to solve 9×9 SUDOKU puzzles automatically. To this end, a captured puzzle image is used, the numbers in this image are recognized by using Artificial Neural Networks (ANN) and a 9×9 number array with these numbers is constituted, respectively. Then, the proposed method is applied to the prepared numerical array for solving the puzzle. The validity of the proposed method is demonstrated with results from an example 9 × 9 SUDOKU puzzle image.*

Keywords: *SUDOKU, Puzzle Solving, Artificial Neural Networks, Image Recognition, Training.*

1. Introduction

Computer games which have various mathematical, algorithmic and visual properties are important research issues of computer science. Hence, they can be used for solving real world problems such as education [1-7], mathematics [8-15] and fitting problems [16-18].

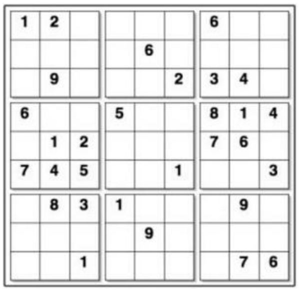
SUDOKU is a puzzle-typed game not related to mathematics directly although it has numerical components. The game is based on filling a 9×9 grid so that each column, each row and each of the nine 3×3 sub-grid contains all of the numbers from 1 to 9. Examples of standard and modified SUDOKU grids are shown in Fig.1.

There are various studies which examine SUDOKU in the related literature. A hybrid AC3-tabu search algorithm is used for solving the puzzle [14]. Permutations are generated for candidate values of empty cells in SUDOKU puzzle [15]. Some studies analyzed SUDOKU considering it as a subclass of the Latin squares [19, 20] and the rules of the puzzle for solving are rewritten [21].

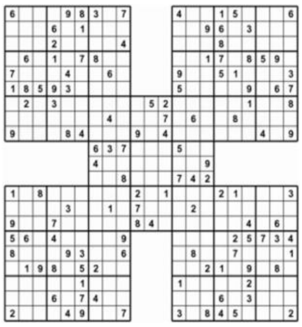
This study focuses on proposing a method for solving a standard SUDOKU puzzle by making use of image processing and ANN. The rest of the paper is organized as follows: in Section 2, the ANN structure is explained; in Section 3, our proposed method is presented and finally in Section 4, the conclusions and future work are given.

2. Artificial Neural Networks

In recent years, image processing has found many application areas such as medicine, engineering, security etc. ANN is an eficient tool used in image processing. The ANN structure and the training algorithm used in this study are back-propagation artificial neural network (BP-ANN) and Levenberg-Marquardt (L-M) algorithm [22, 23].



(a) Standart SUDOKU grid



(b) Modified SUDOKU grid

Figure 1. Example SUDOKU images

The ANN model used in the study consists of three layers of neurons as input, hidden and output as depicted in Fig. 2. The input layer of this system consists of the number images which are represented by matrices. In the training process of this type of network, the connection weights are

Received on: 27.09.2016
Accepted on: 31.01.2017

updated to minimize the error between the correct and estimated values of the system variables [24].

A hidden or output unit in the ANN operates as follows :

$$y_j = f(\sum_i w_{ji} x_i + b_j) \tag{1}$$

where

y_j: transformed output by the jth hidden or output node,

f : activation function,

w_{ji} : the synaptic weight from the ith node to jth node,

x_i : input node,

b_j : bias at jth node

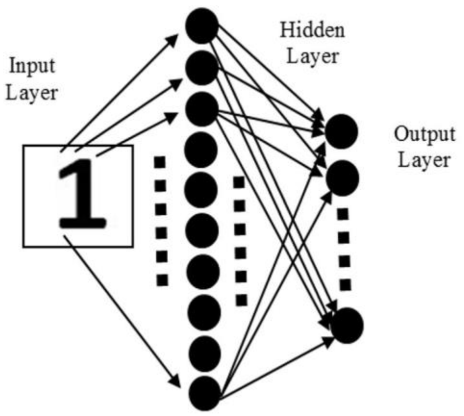


Figure 2. ANN structure used in this study

3. Proposed Method

3.1. Image Recognition

In this section, the steps for image recognition as image segmentation, system training and numerical array constitution are explained. The resolution of input SUDOKU image is 532 × 474 pixels.

Image segmentation: The aim of this step is dividing a one-piece grid image into 81 cells (sub-images). The cells have the possibility of containing 1-9 numbers or they can be empty. Sub-images are transformed to binary images, then, edge detection process is applied to determine whether there is an object or not. If any, the numerical value of the object is recognized by ANN.

System training: In the ANN, there are 3135 inputs representing 55 × 57 pixels in every image of the training set, 9 outputs (1-9 numbers) and a hidden layer with 10 neurons. The input image is transformed to a binary image, then, the matrices for numbers are transformed to a column vector form. The 9 column vectors are collected in an input array. As output value, a 9 × 9 identity matrix is constituted. The input values are used as 70%, 15% and 15% for training, validation and test, respectively. The error value is chosen as 10^{-7}. The 9 output neurons produce outputs which must be 0 or 1. The value of 1 in the column represents the desired number.

Numerical array constitution: The sub-images in the study are transformed to a 9 × 9 numerical array via the principles below:

- if there is a number in the sub-image, this determined number value is placed to the corresponding index in the array,
- if any number cannot be determined in the sub-image, the corresponding value in the array is 0,

otherwise -1 value is placed to the array.

3.2. Puzzle Solving

The algorithm must firstly decide if the current element of the array is 0 or not. Since 0 value means it is an empty cell in the SUDOKU and the appropriate value must be replaced to corresponding element, a candidate vector as ([1 2 3 4 5 6 7 8 9]) is constituted. If the element is not 0, a [X 0 0 0 0 0 0 0 0] vector is constituted, where X is the numerical value of the element. The whole algorithm is depicted in Fig. 3.

The explanation of the proposed algorithm is given below:

A[i][j]: The matrix which has all the recognized numbers in the SUDOKU grid.

Temp A[i][j][ ]: The temporary array which has vectors as [1 2 3 4 5 6 7 8 9] or [X 0 0 0 0 0 0 0 0] instead of each element in A[i][j] according to being 0 or not.

CANDIDATE: The [1 2 3 4 5 6 7 8 9] vector which consists of all the possible choices.

STEP 1 : For all i and j, check whether A[i][j] is 0 or not.

If so,

Temp A[i][j][ ] = [1 2 3 4 5 6 7 8 9] (it means, the appropriate value of the position will be searched from this row)

Otherwise

Temp A[i][j][ ]=[[A[i][j]] 0 0 0 0 0 0 0 0] (it means the appropriate value is settled to the position)

STEP 2: Eliminate the known values in the ith row of Temp A from the CANDIDATE vector.

STEP 3: Eliminate the known values in the jth column Temp A from the CANDIDATE vector.

STEP 4: Divide Temp A array into 9 3 × 3 sub-arrays.

STEP 5: Eliminate the known values in all 3 × 3 sub-arrays in Temp A from CANDIDATE vector.

STEP 6: Put the 3×3 sub-arrays together as Temp A[i][j][ ] array again.

STEP 7: If any Temp A[i][j][ ] element from CANDIDATE vector ≠ 0 GOTO Step 1

Else it means the puzzle is solved properly.

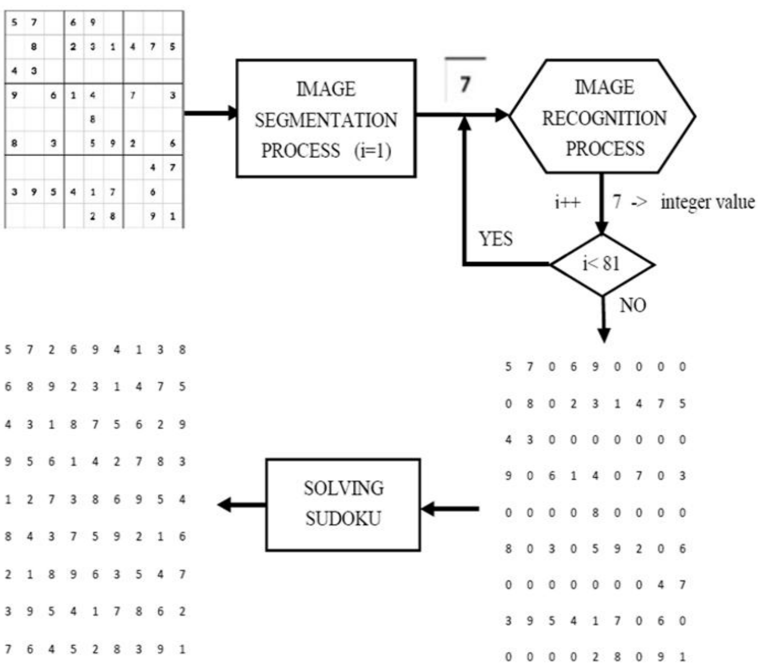


Figure 3. The flowchart of the proposed algorithm

3.3. Experimental Results

The captured 9 × 9 SUDOKU image of 32 × 474 pixels is given in Fig. 4. By the segmentation of that image, 81 sub-images (with or without a number) are obtained. Example sub-images obtained from this process are given in Fig. 5. Next step includes the transformations between RGB form and binary form of the image. Additionally, edge detection and number recognition are also implemented in this step. For instance, the recognition steps of number 5 are depicted in Fig 6.

The number recognition step of this study in the Fig. 6 can be also called training procedure. We use the number set for the training process which is shown in Fig. 7. The fonts of the number set in SUDOKU image and in ANN training procedure differ from each other for the purpose of providing the independency of fonts. An example of ANN training results is shown in Fig. 8. As it can be seen easily from the figure, the system reached the desired error value in 1613th iteration.

The recognition results of numbers 1, 2 and 5 are shown in Fig. 9 as an example. In each column, the row whose numerical value is closest to 1 represents the desired number. Then, the recognized numbers constitute an array as explained in Section 3.1 and this array is given in Fig.10. As explained in Section 3.2, the puzzle solving algorithm is implemented to the array. The temporary array (Temp A in the algorithm) is constituted as in Fig. 11.

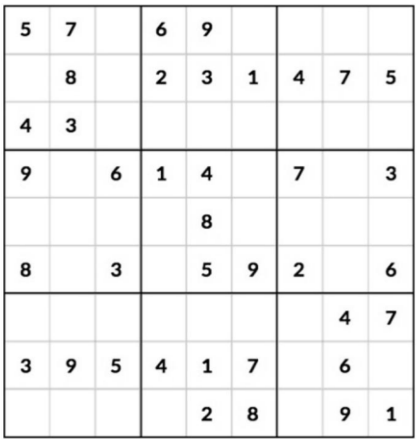


Figure 4. The SUDOKU image of 532 × 474 pixels

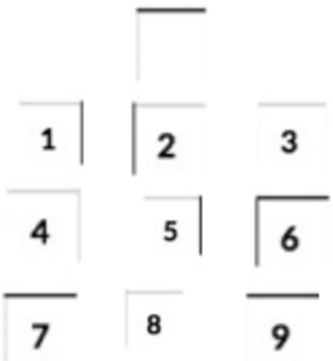


Figure 5. An example of sub-images obtained by using segmentation process

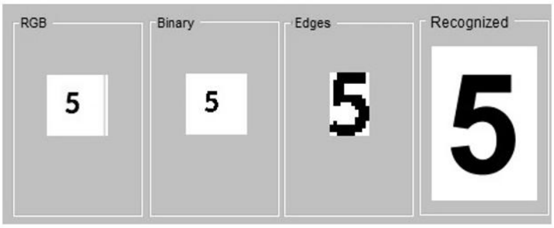


Figure 6. Image processing steps of number 5

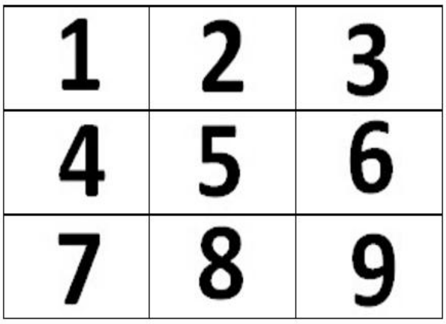


Figure 7. The font of ANN training number set

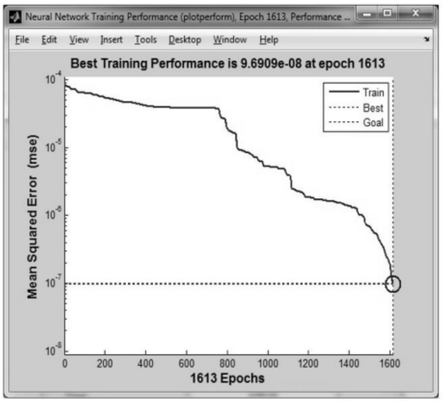


Figure 8. An example of ANN training result

x =	x =	x =
0.9555	0.0004	0.2927
0.0000	0.9827	0.0000
0.0000	0.0026	0.0000
0.0031	0.0019	0.1344
0.0366	0.0000	0.7936
0.0000	0.0000	0.0000
0.0037	0.0007	0.0014
0.0000	0.0007	0.0000
0.0000	0.0000	0.0000

Figure 9. The recognition results of numbers 1, 2 and 5

5	7	0	6	9	0	0	0	0
0	8	0	2	3	1	4	7	5
4	3	0	0	0	0	0	0	0
9	0	6	1	4	0	7	0	3
0	0	0	0	8	0	0	0	0
8	0	3	0	5	9	2	0	6
0	0	0	0	0	0	0	4	7
3	9	5	4	1	7	0	6	0
0	0	0	0	2	8	0	9	1

Figure 10. The numerical array

500000000	700000000	123456789	600000000	900000000	123456789	123456789	123456789	123456789
123456789	800000000	123456789	200000000	300000000	100000000	400000000	700000000	500000000
400000000	300000000	123456789	123456789	123456789	123456789	123456789	123456789	123456789
900000000	123456789	600000000	100000000	400000000	123456789	700000000	123456789	300000000
123456789	123456789	123456789	123456789	800000000	123456789	123456789	123456789	123456789
800000000	123456789	300000000	123456789	500000000	900000000	200000000	123456789	600000000
123456789	123456789	123456789	123456789	123456789	123456789	123456789	400000000	700000000
300000000	900000000	500000000	400000000	100000000	700000000	123456789	600000000	123456789
123456789	123456789	123456789	123456789	200000000	800000000	123456789	900000000	100000000

Figure 11. Temporary array

In the next step, the rows of temporary array are checked and the known numbers are eliminated since they cannot be a candidate for the solution. The new version of the temporary array is given below (Fig. 12). The elimination process is implemented to columns similar to rows. Fig. 13 shows the new version of the temporary array. The elimination process is implemented to 3 × 3 sub-arrays similar to rows and columns. In Fig. 14, the final version of the temporary array is shown. These elimination steps must be repeated until each CANDIDATE vector has a single value. The final solution of the puzzle in this study is shown in the Fig. 15.

4. Conclusion and Future Work

In this study, we have implemented a hybrid SUDOKU puzzle solving algorithm for the purpose of recognizing the numbers in a SUDOKU image and finding the solution of the puzzle. Our study differs from similar studies in the literature via the reasons below:

- it considers both number images and empty square images in the same way,
- it transforms all images (numbers and empty cells) to a numerical array.

We have also observed that the resolution of the image, the noise in the image and font of the texts have an important effect on the performance of the proposed algorithm. We should mention here that our current paper is an extensively improved version of [25].

In future work, the proposed algorithm may be improved by using new images and image recognition methods. Also, it is known that some computer games such as Tetris and SOKOBAN are used for real-world fitting problems. Therefore, we think that there is a possibility for using SUDOKU for the same purpose. This study constitutes the first step of our thought and this algorithm may be improved for real-world fitting problems.

5	7	12348	6	9	12348	12348	12348	12348
69	8	69	2	3	1	4	7	5
4	3	1256789	1256789	1256789	1256789	1256789	1256789	1256789
9	258	6	1	4	258	7	258	3
12345679	12345679	12345679	12345679	8	12345679	12345679	12345679	12345679
8	147	3	147	5	9	2	147	6
1235689	1235689	1235689	1235689	1235689	1235689	1235689	4	7
3	9	5	4	1	7	28	6	28
34567	34567	34567	34567	2	8	34567	9	1

Figure 12. The new version of the temporary array after row elimination

7	1248	6	9	234	13	1238	248
8	9	2	3	1	4	7	5
3	12789	5789	67	256	1569	1258	289
25	6	1	4	25	7	258	3
12456	12479	3579	8	23456	13569	1235	249
14	3	7	5	9	2	1	6
1256	1289	3589	6	2356	13569	4	7
9	5	4	1	7	8	6	28
456	47	357	2	8	356	9	1

Figure 13. The new form of the temporary array after column elimination

5	7	12	6	9	4	13	1238	28
6	8	9	2	3	1	4	7	5
4	3	129	578	7	5	169	1258	289
9	25	6	1	4	2	7	58	3
127	1245	1247	37	8	236	59	15	49
8	14	3	7	5	9	2	1	6
126	126	128	359	6	356	35	4	7
3	9	5	4	1	7	8	6	28
67	46	47	35	2	8	35	9	1

Figure 14. The final version of the temporary array after 3 × 3 sub-array elimination

5	7	2	6	9	4	1	3	8
6	8	9	2	3	1	4	7	5
4	3	1	8	7	5	6	2	9
9	5	6	1	4	2	7	8	3
1	2	7	3	8	6	9	5	4
8	4	3	7	5	9	2	1	6
2	1	8	9	6	3	5	4	7
3	9	5	4	1	7	8	6	2
7	6	4	5	2	8	3	9	1

Figure 15. The final solution

Acknowledgment

The authors wish to thank for the support of the Research Fund of Istanbul University under Project N-53905.

5. References

[1] M. Simkova, "Using of Computer Games in Supporting Education", *Procedia - Social and Behavioral Sciences*, 141, pp.1224-1227, 2014.

[2] Y.G. Butler, "The use of computer games as foreign language learning tasks for digital natives", *System*, 54, pp.91-102, 2015.

[3] G.G. Smith, M. Li, J. Drobisz, H.R. Park, D. Kim, and S.D. Smith, "Play games or study? Computer games in eBooks to learn English vocabulary", *Computers and Education*, 69, pp.274-286, 2013.

[4] N. Alias, F. Rosman, M.N.A. Rahman and D. Dewitt, "The Potential of Video Game in Malay Language Learning for Foreign Students in a Public Higher Education Institution", *Procedia - Social and Behavioral Sciences*, 176, pp.1020-1027, 2015.

[5] I. Ahmad and A. Jaafar, "Computer Games: Implementation into Teaching and Learning", *Procedia - Social and Behavioral Sciences*, 59, pp.515-519, 2012.

[6] H.D. Mo and R.G. Xu, "Sudoku Square - a New Design in Field", *Acta Agronomica Sinica*, 34(9), pp.1489-1493, 2008.

[7] R. Kampf and E. Cuhadar, "Do computer games enhance learning about conflicts? A cross-national inquiry into proximate and distant scenarios in Global Conflicts", *Computers in Human Behavior*, 52(C), pp.541-549, 2015.

[8] H. Mahmoudi, M. Koushafar, J.A. Saribagloo, and G. Pashavi, "The Efect of Computer Games on Speed, Attention and Consistency of Learning Mathematics among Students", *Procedia - Social and Behavioral Sciences*, 176, pp.419-424, 2015.

[9] M. Bakker, M. Heuvel-Panhuizen and A. Robitzsch, "Effects of playing mathematics computer games on primary school students multiplicative reasoning ability", *Contemporary Educational Psychology*, 40, pp.55-71, 2015.

[10] F. Ke, "Computer-game- based tutoring of mathematics", *Computers and Education*, 60(1), pp.448-457, 2013.

[11] F. Ke, "An implementation of design-based learning through creating educational computer games: A case study on mathematics learning during design and

computing", *Computers and Education*, 73, pp.26-39, 2014.
[12] J. Cooper and A. Kirkpatrick, "Critical sets for Sudoku and general graph colorings", *Discrete Mathematics*, 315-316, pp.112-119, 2014.
[13] S.K. Jones, S. Perkins and P.A. Roach, "Properties, isomorphismsand enumeration of 2-Quasi- Magic Sudoku grids", *Discrete Mathematics*, 311(13), pp.1098-1110, 2011.
[14] R. Soto, B. Crawford, C. Galleguillos, E. Monfroy and F. Paredes, "A hybrid AC3-tabu search algorithm for solving Sudoku puzzles", *Expert Systems with Applications*, 40(15), pp.5817-5821, 2013.
[15] A.K. Maji, S. Jana and R.K. Pal, "An Algorithm for Generating only Desired Permutations for Solving Sudoku Puzzle", *Procedia Technology*, 10, pp.392-399, 2013.
[16] D. Dor and U. Zwick, "SOKOBAN and other motion planning problems. *Computational Geometry*, 13(4), pp.215-228, 1999.
[17] R.A. Hearn and E.D. Demaine, "PSPACE-completeness of sliding-block puzzles and other problems through the nondeterministic constraintlogic model of computation", *Theoretical Computer Science*, 343(1-2), pp.72-96, 2005.
[18] T. Anthony, D. Polani and C.L. Nehaniv, "General Self-Motivation and Strategy Identification: Case Studies Based on Sokoban and Pac-Man", *IEEE Transactions on Computational Intelligence and AI in Games*, 6(1), pp.1-17, 2014.
[19] M. Mahdian and E.S. Mahmoodian, "Sudoku Rectangle Completion", *Electronic Notes in Discrete Mathematics*, 49, pp.747-755, 2015.
[20] R. Bejar, C. Fernandez, C. Mateu and M. Valls, "The Sudoku completion problem with rectangular hole pattern is NP-complete", *Discrete Mathematics*, 312, pp.3306-3315, 2012.
[21] G. Santos-Garcia and M. Palomino, "Solving Sudoku Puzzles with Rewriting Rules", *Electronic Notes in Theoretical Computer Science*, 176, pp.79-93, 2007.
[22] K. Levenberg, "A method for the solution of certain problems in least squares", *Quarterly of Applied Mathematics*, 5, pp.164-168, 1944.
[23] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters", *SIAM Journal on Applied Mathematics*, 11, pp.431-441, 1963.
[24] S. Haykin, "Neural Networks and Learning Machines" (3rd ed.). Prentice Hall.
[25] S. Sevgen, E. Arslan and R. Samli, "Number Recognition of Sudoku Grid Image with Artificial Neural Networks", *In Proceedings of 22nd International Conference on Neural Information Processing (ICONIP 2015)*, pp.540-547, 2015.

Selcuk Sevgen is currently an Assistant Professor at the Department of Computer Engineering in Istanbul University Istanbul, Turkey. He received his M.Sc. and Ph.D. degree in same department in 2003 and in 2009, respectively. His main interests are Neural Networks, CNNs.

Emel Arslan was born in Istanbul, Turkey, in 1977. She received the B.Sc. and M.Sc. degrees from Trakya University, Edirne, Turkey, and Ph.D. degree from Istanbul University, Istanbul, Turkey, in 2001, 2004 and 2011, respectively.
She is currently working as an assistant professor in the Department of Computer Engineering, Istanbul University.
Her research interests are artificial neural networks, natural language processing, image processing applications and intelligent systems..

Ruya Samli is currently a Associative Professor at the Department of Computer Engineering in Istanbul University, Istanbul, Turkey. She received her M.Sc. and Ph.D at the same department in 2006 and 2011, respectively about stability of different types of neural networks. Her main interests are Neural Networks and modelling techniques.

SPIT DETECTION AND PREVENTION

Selin KAMAS[1], Muhammed Ali AYDIN [2]

[1]Netaş Telecommunication, Istanbul, Turkey
[2]Department of Computer Engineering Istanbul University, Istanbul, Turkey
skamas@netas.com.tr, aydinali@istanbul.edu.tr

Abstract: *In telecommunication technology VoIP protocol has become a very popular technology as it is cheap, efficient. Also it has easy deployment. While it has lots of advantages it brings lots of vulnerabilities. These are Man in the middle Attack, Replay Attack, Teardown Attacks, Flooding Attacks, Toll Fraud and SPIT (Spam over IP Telephony). Spam over IP Telephony (SPIT) is an known threat in the Voice over IP Networks (VoIP). Even though evolved from email spam, SPIT is more obstructive and intrusive in nature. SPIT attack is called important threat of reliability and availability of VoIP system and also it is difficult to make SPIT call in PSTN (Public Switched Telephone Network) system. In this work It is tried to say how SPIT attacks occur, how attackers do it and also it is mentioned that prevention mechanisms and compare them in terms of feasibility, advantages and disadvantages..*
Keywords: *VoIP, VoIP security, SPAM over IP Telephony ,SPIT,Captcha, Whitelist, Blacklist.*

1. Introduction

VoIP spam is unwanted and automatic calls that are consecutive records have been recorded previously. VoIP system has much vulnerability because of its IP Infrastructure. One of them is SPIT. In early 2004 found 50% of e-mail is determined to be spam call. E-mail and old phone system protocol's addressing system is similar with VoIP. Therefore VoIP system is vulnerable spam call, too. In VoIP systems, Spamming reveals more effective results than can be done in e-mails protocols. Because spam calls obstruct people to use phone. Additionally, VoIP systems are cheaper than PSTN system. Therefore this makes VoIP system an easier target for telemarketers [1]. Telemarketers are people who make unwanted phone calls to sell products or services.

VoIP protocols have a lot of tools (SIPp, Asterisk) that are used by attackers to make spam call[2]. The other reason VoIP system vulnerable to SPIT is SIP (Session Initiation Protocol)'s vulnerabilities. The Session Initiation Protocol (SIP) is a communications protocol for signaling and controlling multimedia communication sessions. SIP characterizes the messages that are sent between endpoints, which govern establishment, termination and other essential elements of a call. The protocol can be used for creating, modifying and terminating sessions consisting of one or several media streams. SIP has information about voice, codec, application type and status of call [3]. "Spammers" starts session and If they use SIP, used request message type is "INVITE" to start session. After called answer call, they send automated voice record (SIPp, Asterisk) to spam called. These calls are unwanted, irrelevant, unsolicited and unexpected and called SPIT. To make SPIT call is ordinary and its result is effectively dangerous. Therefore, to make secure VoIP system, providers should apply variety of methods to detect and prevent SPIT call. To mention SPIT attack's visibility and its effective dangerous result Softbank in Japanese reported they have seen three big SPIT attack in their VoIP system [15]. Voice over IP (VoIP) is a methodology and group of technologies for the distributing of voice communications and multimedia sessions over Internet Protocol (IP) networks, such as the Internet. To prevent SPIT call; it is not idea to change its IP infrastructure. Using IP protocol has common usage and not opens the change [5].

2. Compare SPIT and SPAM

SPAM e-mails do not disturb users or system until they open their e-mails. After they open e-mails they can understand this is spam mail and they can point this mail's sender as a spammer. E-mail protocol can block this sender to prevent send spam mail after this point. Also spam mail can detect before user open mail by checking content of mail. But SPIT call cannot prevent like that scenario. Because, SPIT calls disturb user when users open call. And until users open the call system cannot understand it is SPIT because content of

Received on: 10.06.2016
Accepted on: 08.02.2017

voice communication cannot be seen. And this prevents user access to service.

Also VoIP systems and e-mail protocols have different in terms of time. VoIP systems work in real time.

Because of e-mail service is content-based, spam mails can be detect by checking content but in VoIP system cannot check content of communication until conversation starts. Moreover, filtering cannot be made by looking content of conversation.

The following table shows the system in PSTN or VoIP systems is that when compared to the costs of spamming.

Table 1. Comparison of Costs of Spam Attacks on PSTN and VoIP systems [8]

Cost	SPAM (PSTN)	SPIT (VoIP)	Note
Software Cost.	A.	A.	**A** (change according to signaling protocol).
Hardware Cost.	10B-100B.	B.	**B** (does not change according to signaling protocol).
Cost of every spam.	About 1000C.	C.	**C** (does not change according to signaling protocol).

3. Spamming Over Internet Telephony (SPIT)

Attackers make phone unwanted and unexpected continuous call to inhibit users to access services or to advertise or discredit providers[6]. VoIP system vulnerable this attack the same reason with e mail services. This reason is every person can call every person really cheaply.

As just mentioned, Telemarketers also benefit from SIP addressing that like email addressing. Telemarketers use several web pages, e-mail lists or crawling technic to takeover SIP addresses. Also to seizure SIP addresses or usernames attackers make Brute Force or Dictionary Attack.

While attackers make SPIT , for example If there is and 30sn packet to send, attackers use RTP (Real Time Protocol) and it takes 30sn to deliver this packet and a system security administrator can think this feature can be used to prevent SPIT. But it is not idea because telemarketers use parallelism to handle this condition [5].

Voicemail services are services that facilitate the feasibility of SPIT attacks. Telemarketer send SPIT call even offline user thanks to previously recorded messages.

SIP provide user to be anonym.. SIP enables this capability through e-mail services, unlike VoIP protocol circuit-switched system, resulting in the vulnerability of a spamming attack. SPITTers create a botnet for themselves and hide their IP addresses. SPITTers are the same people with telemarketers. A botnet (also known as a zombie army) is a number of Internet computers that, although their owners are unaware of it, have been set up to forward transmissions (including spam or viruses) to other computers on the Internet.

In one experiment, without any SPIT attack network usage is 21kbps and for 30sn needed space is 75KB and in 30sn 100 spam e-mail can be send in experimental network. From this point 100.000 voice record and every one lasts 30sn. Therefore it shows that this SPIT calls needed 7.2GB uplink network usage capacity. From there SPIT calls can be detected [5].

List of requirement of architecture to prevent SPIT

- ✓ Do not block legal users
- ✓ Maximize possibility of detect attacker who make SPIT calls
- ✓ Stop communication with attacker and victim called
- ✓ Prevent the SPITTers to define themselves as legal
- ✓ Be used as appropriate for different language, infrastructure, environment (office, home) [8].

To prevent SPIT there are a lot of methods but none of them have all of requirement which are mentioned above. The methods should exclude called users while detect or prevent SPIT calls. Based on this assumption; get feedback from a caller will be way more intelligent solution. From this point there is an algorithm which defines black and grey list. While prevention mechanism make classification, it check caller from inter-domain. Secondly, mechanism have waited proof from caller about his*her honesty about call goal's. This proof can be done with Computational Puzzles, sender checks, Turing test etc. But with this kind of mechanism problem is that: caller has to proof his/her honesty and this cause users wait long time until call establish. Moreover, computational puzzles and Turing test's complexity is not effective to implement real time application even though their complexity is median [8].

3.1 Solution Methods

In e-mail services if users can manage their e-mail individually, defining black and white list approaches to prevent spam mail in level of proxy and client can be feasible. Because users should able to edit their mail according to type of mail lists (spam, social, advertisement, all etc.).However service providers should able to filter e-mail in their servers. Also defining black and white list merely is not enough to prevent spam mails because of ability of create botnet or IP spoofing attack.

Another method is CAPTCHA (Completely Automated Public Turing to Tell Computers and Humans Apart). In many places, end users expected from brute-force authentication mechanism used to verify user to prevent attacks on the proxy. This verification mechanism's random code instantaneous transmission can be produce in Proxy (on the fly)'s process or using sound recordings produced by the user [8]. But expecting user to strive in this protection mechanism is not very accurate. In addition, the reliability of the records carried out on-the-fly process should also be discussed.

Transmission of these records must be secured using some encryption methods. This requires a distinct performance.

Other method is non-reputation. From historical call details can be found caller and called information. But SPITTers have found ways to overcome this. They have agreed with peer and pretending as a legal user. After that they start to SPIT call and handle non-reputation mechanism. Non-reputation does not require any effort to SPITTers is a deterrent method.

To define White list is somehow limits SPITTers. But defining black list is not efficient method to limit or stop SPITTers to make SPIT call as it is mentioned before. Reason of this is ability to create botnet or dynamic IP addresses etc. [8].

Using CAPTCHA in web pages as a Picture or text is common way. Even the use of the Web page has security vulnerabilities. It should not be defined directly in the codes. Hash algorithms should be used during displaying of these numbers to users. This method cannot be used in e-mail services because e-mail services works asynchronous. It can be used in voice transmission but this prevention mechanism should be secure, too. Users should not have to expend extra effort for this mechanism [8].

Other prevention mechanism called Domain Based Authentication and Policy Enforced for SIP(DAPES) and it uses TLS(Transport Layer Security) and digest authentication mechanism[11]. Implementation of this method is infeasible and complicated. Because implementing this method require to change other modules.

In RFC5039; there are lots of methods to prevent SPIT. These are content filtering, black and white list, Consent-Based Communications, Reputation Systems, Address Obfuscation, Limited-Use Addresses, Turing Tests, Computational Puzzles, Payments at Risk, Legal Action, Circles of Trust. All of these have some disadvantages and because of this it is mentioned that every of methods has some vulnerabilities [12].

1. Content Filtering

Content filtering is a method which is used in e-mail services to prevent spam mails. However in VoIP system cannot be applied because voice communication is real time and no one check content until called answer to call. If content is saved as a voice record in voice mail, this method can be applied. In this case, to control this content, prevention mechanism should have sound/video recognition algorithm. But these algorithms can be broken by attackers and also these algorithms are complex and hard to implement. In addition in sound recognition system %40 of sound is noise [14].

2. Black List and White List

Black list is not a best practice for SPIT attack detection a prevention mechanism even though in e-mail services. Although SIP protocol makes inter-domain authentication, attackers can create limitless addresses and they do not care of being in black list.

White list and black list work oppositely. Attackers want to be white list. If SIP authentication mechanism work truly, SPIT calls will be detected. If there is a black list protection, users who are not in white list make call with effort. This is a restrictive method in terms of users comfort. Also expecting users to be in white in in their first call is not expected way. Moreover first SPIT calls will not be detected [17]. Also looking universal list and try to detect SPIT call is not pragmatic or feasible way to protecting from SPIT calls.

3. Grey List

Defining grey list; when users make their first call they will be in grey list and after a while system want to users to make call again. If user make call in this specific time line, he/she will be in white list else black list [17]. This is more feasible approach is based on previous black / white list identification.

D.Shin[17] define two grey list identification according to duration time of calls. There is a threshold for duration of call.

If duration of call longer then this threshold, call be marked as a SPIT call. Authentication mechanism of this method is weak. Therefore is not applicable [17].

4. Consent-Based Communications

This is hybrid solution of black/white list protection approaches. Users can accept calls directly or request for authentication. At first glance it seems applicable but its authentication mechanism is not sufficient for detect/prevent SPIT calls in VoIP systems [12]

5. Reputation Systems

This method is also hybrid soltion of black/white lists. For example, If A user is not white list for B user, non-reputation system helps B user to accept or reject this call. Non-Reputation system is used in more central messaging architecture. Non-reputation score calculates from user's feedback and according to this result system decide about call's intention. From this point, there will be same problems with black list. Because there will be generally positive feedbacks.

6. Address Obfuscation

SPITTers generally find e-mail addresses of SIP users from web pages or public places. In these circumstances, e-mail addresses should be hidden and should be too complex to be non-predicted.
Address Obfuscation is an approach for this situation. It advice that while saving e-mail addresses, these should be formatted differently as a non-predicted. For example, user@domain.com e-mail address should be saved as "user at domain com" or "j d r o s e n a t e x a m p l e d o t c o m".
However, under these conditions, after attackers notice pattern of format, can create a tree and it is possible to turn around [12].

7. Limited-Use Addresses

Limited-Use Addresses is about address obfuscation method. It limits number of user's e-mail addresses. For example, the number of e-mail addresses of specific users can be limited within specific time-line
After time-line user's access of e-mail address should be denied. If in this method, user's current e-mail address is used to make SPIT call, protection mechanism works and after that time this e-mail address cannot be used.
A disadvantage of this method is if user's e-mail address reaches maximum user has to notify other users who will be called from this email address. It is an expectation that this will not be welcome by users [12].

8. Payments at Risk

With this approach, for example If A user calls B user, Firstly A user needs to pay for a call to B user. If B user voted this call as a normal call, payment of this call repay to A user. Disadvantages of this method there is an need of transition payment two times and If A user do not have enough money to pay, even if A user is a normal user, he/she will not call B user.

9. Model-Based Filtering

It creates a model based on actual calls over. When the model was created, the frequency and duration of calls are compared with the previous calls. Also while creating a model user based call number, repetitive number of calls, time of calls, and number of unknown caller are saved. After calculation these numbers, decision of forwarding calls to called user depends on these results. This method is also not effective and performance. There can be high false positive rate with just these metrics [14].

10. Circles of Trust

With this approach users voted caller about call's intention. Trust score is applied about joining or rejecting from conversation. Calculating trust score is user-based because of this reason applying this method is proper small network or small providers. Applicability of this method in big networks is not enough to provide scalability.

11. Dendritic Cell Algorithm

DCA (Dendritic Cell Algorithm) is abstract model based classification method and it uses dendritic cell methodology in biology as a prevention mechanism.
Dendritic cells' task is antigen in biology [16]. DC provides to detect bacteria, viruses and other parasites in body. In term of performance and accuracy DC algorithm gives effective results and it is commonly used to detect and prevent for network security [17]. DCs are used as a key for detection and prevention algorithm for SPIT calls.
DCA works real time and is used to detect the anomaly on the data time-series. It is processing the signals and inform about the status of the network. DCA algorithm divide signal into four type [17]. These are;

- **PAMP Signal;** PAMP (Pathogenic Associated Molecular Pattern) signals generate micro-organism. Therefore, if there is PAMP Signal in network, it shows there is a high level anomaly in system.
- **Danger Signal;** This signal Show sudden death in biology. As a detection mechanism If there is an danger signal in the network If show anomaly but smaller than PAMP signal.
- **Safe Signal;** It shows that network is secure.

Antigen signs DC have changed somehow and DC's states can be three type. These are **immature**, **semi-mature** and **mature.**
It is shows transition within states of DC's in Figure 1.

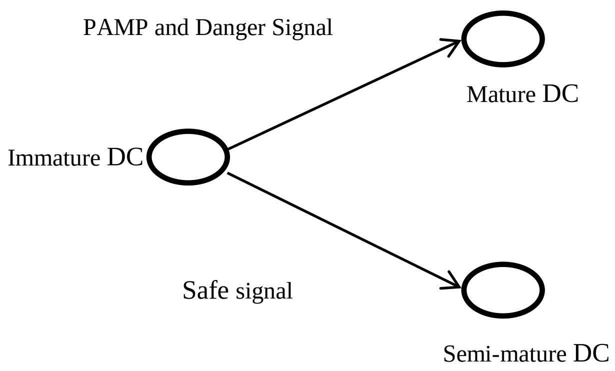


Figure 1: DC State Transition[17]

Inflammation: It shows immature DC have not passed mature state in biology, yet. In anomaly detection, the signals are effective in the formation of the other three signals.
If immature DC receives PAMP or danger signals, it switches to mature DC. Otherwise switches semi-mature.
DC algorithm gives three outputs. These are;

- **CSM Output Signal;** Costimulatory Molecule (CSM) signal show threshold of mature and immature DC. Before moving to the lymph node it is location of the incoming signal.
- **Semi-mature Signal;** It shows the cumulative sum of the safe signals.
- **Mature Signal;** It shows the cumulative sum of the PAMP and danger signals.

DC's states can be shown by 0(semi-mature) or 1(mature). DCA algorithm uses the time difference between the last and first call, daily call numbers to calculate **PAMP** signals, failed call numbers and time duration of calls are calculated for **danger** signal and lastly number of established/successful call numbers is calculated for **safe** signal. Using these numbers CSM and state of DC are determined [17].
After using DCA algorithm for detection and prevention SPIT calls, test results give %93.33 accuracy rates to SPIT calls and with % 96.67 accuracy rate, normal calls are classified. Even though DCA algorithm is complex and costly it can be implemented. And also DCA algorithm can be used to detect other anomaly of the system (flooding, DoS/TDoS, fuzzing, malformed SIP message etc.) not just for SPIT calls. In DCA algorithm to reduce false negative/false positive possibility, there can be defined more metric to calculate state of DC and CSM output. It will be increase accuracy of algorithm. For example, difference between normal and SPIT calls, ID of SPIT calls generally calls someone, not be called too much. ID of SPIT call's incoming call rate is much less.

If there is a big difference between incoming and outgoing call, can be sign PAMP signal. In addition, from historical data, SPITTers generally does not call same number again. However normal users generally call the same number more. The difference between repetitive and different call rate can be used PAMP and danger signal. It can reduce false negative/positive rates By using these metric additionally.

4. Conclusions

SPIT attacks are a threat for VoIP users and infrastructure. There are many method for detection and prevention and as a mentioned above also all of methods have advantages and disadvantages. Therefor there should be combined some of solution approaches. For example in DCA algorithm metric numbers should be increased and there can be used some machine learning algorithm to learn model of system. They can use support vector machines to classify calls [20]. Moreover they can use neural network algorithm and other machine learning algorithm for classification. Additionally they should define distinctive feature [+38] of SIP and use them as a feature vector. After classification traffic can be classify as a normal and bad. After obtained test data realizing SPIT call will be more accurate. And also with this approach users do not need to effort.

5. References

[1] J. Rosenberg "The Session Initiation Protocol (SIP) and Spam," draftietf-
sipping-spam-01.txt, July 2005. Work in progress.
[2] A. Baxter. Shtoom. http://divmod.org/
projects/shtoom, April 2006.
[3] Snoeren, Alex C., and Hari Balakrishnan. "An end-to-end approach to host mobility." Proceedings of the 6th annual international conference on Mobile computing and networking. ACM, 2000.
[4] L. von Ahn, M. Blum, N. J. Hopper, and J. Langford. The Captcha Project. http://www.captcha.
net/, April 2006.
[5] Bila, Nilton. "Dealing with SPAM in Voice over IP." change 34.4: 37.
[6] FELTEN, E. (2005, March 15). Unwanted Calls and Spam on VoIP. Retrieved from https://freedom-to-tinker.com/blog/felten/unwanted-calls-and-spam-voip/
[7] C. Garretson. Qovia ready to take on VoIP spam.
http://www.networkworld.com/news/
2004/071204qovia.html, July 2004.
[8] Prevention of Spam over IP Telephony (SPIT)Juergen QUITTEK, Saverio NICCOLINI, Sandra TARTARELLI, Roman SCHLEGEL
[9] Edelson, E.: Voice over IP: security pitfalls. Network Security, vol. 2005, pp. 4–7 (2005)
[10] Performance analysis of VoIP spoofing attacks using classification algorithms (2014 Application and Innovation in Mobile Computing)
[11] K. Srivastava, H. Schulzrinne, "Preventing Spam for SIP-based Instant Messages and Sessions", Technical report, Columbia, 2004.
[12] https://tools.ietf.org/html/rfc5039
[13] Secure Layered Architecture for Session Initiation Protocol based on SIPSSO. 2015 12th International Conference on Information Technology - New Generations
[14] A comprehensive SPIT Detection and Prevention framework based on Reputation Model on Call Communication Patterns. Farideh Barghi, Mohammad Hossein, Hossein Khosravi Roshtkhari.
[15] VOIPSA. Confirmed cases of SPIT. Mailing list (2006),http://www.voipsa.org/pipermail/voipsec-
voipsa.org/2006-March/001326.html
[16] The Dendritic Cell Algorithm Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy. Julie Greensmith October 2007
[17] Dendritic Cell Algorithm for preventing Spam over IP Telephony. V.Srihari, P.Kalpana , R.Anitha.2015
[18] D. Shin, J. Ahn, and C. Shim, "Progressive Multi Gray-Leveling: AVoice Spam Protection Algorithm," IEEE Network, vol. 20, pp. 18–24,
September/October 2006.
[19] Detection and Filtering Spam over Internet Telephony-A User-behavior-aware Intermediate-network-based Approach. Yan Bai1, Xiao Su1 and Bharat Bhargava3

[20] NASSAR, M., STATE, R., AND FESTOR, O. Monitoring SIP traffic using support vector machines. In RAID '08: Proceedings of the 11th international symposium on Recent Advances in Intrusion Detection (Berlin, Heidelberg, 2008), Springer-Verlag, pp. 311-330.
[21] On the inefficacy of Euclidean classifiers for detecting self-similar Session Initiation Protocol (SIP) messages. Anil Mehta*, Neda Hantehzadeh*, Vijay K. Gurbanit, Tin Kam Hot, Jun Koshiko* and Ramanarayanan Viswanathan*12th IFIP/IEEE International Symposium on Integrated Network Management 2011

Note:

Selin Kamaş was born in Tunceli. Selin is living in İstanbul/Turkey and is graduated from İzmir Instıtute of Technology. Department of her Computer Engineering(06.2015). Because of her interest of network security, she is working cyber security department of NETAŞ as a Cyber Security R&D Engineer. Her first article is published Toll Fraud Article XX.Internet Conference(2015).

SUPPORT VECTOR MACHINES COMBINED WITH FEATURE SELECTION FOR DIABETES DIAGNOSIS

Fatma PATLAR AKBULUT[1], Aydın AKAN[2]

[1]Department of Biomedical Engineering, Istanbul University, Istanbul, Turkey
[2]Department of Electrical and Electronics Engineering, Istanbul University, Istanbul, Turkey
fatma.akbulut@ogr.iu.edu.tr, akan@istanbul.edu.tr

Abstract: *Clinical Decision Support Systems (CDSS) are used as a service software which provides huge support to clinical decision making process where the main properties of a patient are matched to a tangible clinical knowledge. Within this gathered important information about patients, the medical decisions can be made more accurately. In this paper we present a CDSS that uses four physiological parameters of patients such as Pre-prandial Blood Glucose, Post-prandial Blood Glucose, Hemoglobin A1C (HbA_{1c}) and Glucose in Urine to produce a prediction about the possibility of being diabetic. According to collected reference data provided from hospitals, the disease can be predicted by comparing the input data of patients. If the system cannot procure a prediction about patients' status with these parameters, then the second phase which uses soft computing techniques is put into process with requesting additional data about patients. Our conducted experiments show that the diagnosis can be established in a breeze by getting the patients information with %80 accuracy. Support Vector Machines were applied to achieve maximum success rate with nine different physiological parameters such as; Pregnancy, glucose, blood pressure, skin fold, insulin, Hemoglobin A1C, body mass index, family tree and age. Four different Kernel Functions are implemented in case studies and classification process is optimized by reducing attributes with feature selection algorithms. This represents an improvement in classification of CDSS, while reducing computational complexity.*
Keywords: *Decision Support System, Support Vector Machine, Sequential Forward Search, Feature Selection.*

1. Introduction

Diabetes is a common disease that affects people at all points in our environment [1]. In this section it is simply explained what is diabetes, what kind of things can cause diabetes shall be answered to understand the concept of our Clinical Decision Support System (CDSS). If the patient has high blood sugar, because of inadequate of insulin production, it is apparent that diabetes occurs. In fact, two types of diabetes are considered frequently. Type 1 that is named as insulin-dependent diabetes or early-onset diabetes forms 5-10% of all diabetes cases. In this type of diabetes, the body of human does not produce insulin so that the people who are affected by type 1 of diabetes should take insulin injections for their rest life. Another type of diabetes which is Type 2 causes that the body of human does not produce enough insulin. Approximately 90-95% of people shall encounter this type of a diabetes [2][3]. In addition to these situations, it can be shown that death's risk of women because of the diabetes is higher than man in the world [4]. For our research, Decision Support Systems are suitable to generate solutions about diagnosis of diabetes' types.

Decision Support Systems enable decision maker to designate the alternative solutions and the re-revision of data while trying to solve the problems [5]. In today's world, Healthcare Organizations benefit from Information Systems [6] in the fields of management services, diagnosis of a patient, the decisions to be made about the patients by the doctors, being a guide for both physicians and nurses, interpreting on signals, laboratory services, and patient management and so on. That's why the most preferred system in this field is the Clinical Decision Support Systems that are the computer based systems supporting the physicians and other personnel in the process of clinical decision making [7].

In the proposed system, four physiological parameters of a patient that are Pre-prandial Blood Glucose, Post-prandial Blood Glucose, HbA_{1c} and Glucose in Urine are procured by a hospital in accordance with the official permission and personal rights and personnel information are kept secret. Thanks to these values, the system gives information to the patients whether they are diabetics or not.

Developing a CDSS for a hospital in Istanbul provides some benefits to patients, nurses, physicians so that our goals for the CDSS for Diabetes are;

- To support physicians in order to determine diagnosis of patient data.
- To support physicians in the process of patient's management.
- To be clear in diagnosis of patient data.

Received on: 05.12.2016
Accepted on: 30.12.2016

- To help physicians in their determining diagnosis of patient data and prevent them in order not to make a mistake and if it is possible, to minimize the mistakes as these physicians work under the hard conditions.
- To create a cost cutting atmosphere in the hospitals in their process of diagnosis of patient data.
- To create more reliable physicians for the patients and enhance the relationship between the patients and physicians.

The remaining sections of this paper are organized as follows: section two introduces the background and related work, section three explains the approach in detail, section four shows the details of the experiments, and section five provides the conclusion and future work.

2. Related Work

CDSS is used as a software that provides a huge aid to clinical decision making where the main properties of a patient are matched to a tangible clinical knowledge, then the information are gathered by the clinicians or patients to notify the decision. Furthermore, CDSS decrease medical errors and raise strength and quality of health have become a favorite area in the not too distant past [8].

Below, some further studies handled so far with regards to the subject of Clinical Decision Support Systems have been outlined.

In the study of Kostic et. al. CDSS are used to improve the quality of healthcare delivery. They build a module that is a developed part of KardioNet system whose purpose is providing clinical decision making support for patients' treatment with Acute Coronary Syndrome [9].

El-Fakdi et. al. study is a detailed work that is about the work-flow-based CDSS that aims to give case-specific evaluation to clinicians during the complex surgery. Based on the workflow that was gathered, the software will use a Case-Based Reasoning methodology to get similar past cases from a case base [10].

In the study of Kamaleswaran et. al. thanks to CDSSs, clinicians provide accurate data analysis and recommendations to support health care of patients. They present the process web service as a new method that provides contextual information to a data stream processing CDSS [11].

Kunhimangalam et. al. constructs a system that has 24 input fields that consist of clinical values of diagnostic test and whose output field is the diagnosis of disease whether it is Motor neuropathy, sensory neuropathy, mixed type or normal case. The results were gathered are compared with the clinician's opinion. The system provides clinicians to make prediction of a better diagnosis [12].

Another study is about computerized epilepsy treatment CDSS whose purpose is to aid the clinicians about selecting the best anti-epilepsy treatments. Standridge et. al. has evaluated the system in three areas that are the preferred anti-epilepsy drug choice, the top three recommended choices, and the rank order of the three choices [13].

In the study of King et. al. they build the Genetic Smart Alarm(GSA) that is a framework for the design, analysis, and implementation of CDSS. Their aim was that showing how a GSA can be used to adapt a smart alarm for specific patient populations [14].

Kemppinen et. al. study is CDSS of the diagnosis of ADHD that is a complex neuropsychiatric disorder. The system supports the implementation of the new adult ADHD patient evolution, diagnosis and treatment process [15].

In the study of Karim et. al. they build a virtual telemedicine that uses CDSS is used as a rural station. The system provides diagnosis of patient's disease and it sends an e-mail to clinicians and when the response is received, the CDSS is updated for the future values [16].

The Bayesian Network was built by using systematic response syndrome criteria, mean arterial pressure, and lactate levels for sepsis patients in the study of Gultepe et. al. the resulting network brings to light a clear relationship between lactate levels and sepsis. Bayesian network of sepsis patients hold the show of providing a CDSS in the future [17].

In the study of Mattila et. al. they present a clinical decision support system that takes as an example of a patient's disease situation from heterogeneous multiscale data. Several medical datasets are applied to the system and the system is asserted by implementing a new clinical decision support tool for early diagnosis of Alzheimer's disease [18].

3. Methodology

Our proposed system is implemented as a web application to be used from anywhere in the hospital without making any install processes. Also, doctors can reach to the system from thin clients like PDAs easily.

The reference data is received from Yalova Public Hospital in Turkey. All data is about patients that are treated between the years 2013 and 2014 kept secret, because according to Turkish Rights of Patient Law, 21th clause [19], the information of patients must be kept confidential. Hereinbefore, four parameters Hemoglobin A1C, Pre-prandial Blood Glucose, Past-prandial Blood Glucose and Glucose in Urine are taken from the patients respectively.

Decision support process start with acquiring three main properties of patients as age, gender and health complaints as symptoms. Users may choose any 13 complaints from the system such as; frequent urination, excessive thirst, blurred vision, weakness, fatigue, unexpected weight loss, the feeling of hunger, nausea, vomiting, breath odor, frequency urinary tract infections, dry and itchy skin, slow healing of wounds and total deduction due to their conditions. Possibility of being a diabetic is categorized into 3 levels. First two levels can make a prediction directly without using any soft computing technique. On the contrary, level 3 type patients can be only graded with using Support Vector Machines (SVM). Those predictions are used at the moment of decision.

Our proposed system helps medical crew in order to have necessary information about their patients in a long term. This system consists of sequential stages that are generally compares the inputs that are coming from the patients and values as knowledgebase from the hospital called reference values. The methodology of main flow and the architecture of entire operations are indicated in Algorithm 1.

Algorithm 1: Pseudo code for disease prediction

```
Input: n = Symptom Quantity, k=Cluster Quantity, a = Patient Age, g = Patient Gender, S⃗n = Symptom Vector, W⃗n = Symptom weights Vector, C⃗k = Symptom Cluster Vector,

Output: Level of Disease ← ('Healthy', 'Potential Diabetes', 'Possible Diabetes')

Initialize:
   i:= 1….n, lim←loadUplerLowerLimits(),
f←getTestData(), Threshold:=3

for i=1: n
     W⃗n ← f(n)/ sum(f(n))
end for
α ← CalcFirstDiagnoseRef (n,a,g, W⃗)
if α < Threshold
   Print "Healthy"
end if
else if α ≥ Threshold
    result ← Diagnose(lim)
end else if
```

Weight of the symptoms within the given model is calculated by dividing the number of symptom frequency corresponding to the reference data into the total number of symptom frequency, when calculating the symptom's disease probabilities.

CDSS for Diabetes is used for the diagnosis of the diabetes and the information about the diagnosis is calculated by using reference data. The system works with some comparison to achieve the results and the gaps of comparisons are common information that can be found in some researches [20].

When the system is investigated in depth, the patient's information such as name, surname, age, gender is taken. According to gender some symptoms are shown in the startup screen of the system. Users asked to check these symptoms to evaluate their conditions. If the result that means the symptoms from the patients is less than our threshold (3), it means that person is healthy.

Potential patients are grouped under 7 categories. If the user cannot match one of these categories, then the classifier produces a prediction about patient's medical condition.

However, if the frequency of the symptoms is greater than or equal to 3, patient's pre-prandial glucose is requested, and the system is navigated to Pre-prandial stage. System takes the pre-prandial blood, and then compare with reference function. If the pre-prandial blood glucose value is smaller than 110, the results are compared with reference values. If the given value is less than the reference value, it means that the person is healthy. On the other hand, if the given value is greater than this reference value, patient's post-prandial glucose value is requested as additional information. Another condition is that if the pre-prandial value is greater than 110 but less than 126, post-prandial blood glucose value is requested and average blood sugar is calculated. The results are compared to reference values and according to results, the amount of glucose in the urine of patients is requested.

Algorithm 2: Pseudo code for Diagnose

```
Initialize:
r ← GetPrePrandialBG()
if r < 110 /*Case 1*/
   if r ≥ lim
     r ← GetPostPrandialBG()
           if r < 140
             Print "Healthy"
           else
              Print "Possible Diabetes"
           end else
           end if
     else Print "Healthy"
     end else
     end if
 end if
 if r > 126 /*Case 2*/
      r ← GetPostPrandialBG ()
      if 140 < r < 200
         Print "Potential Diabetes"
      else
         if r < lim
            Print "Potential Diabetes"
         else
            Print "Possible Diabetes"
         end else
         end if
      end else
      end if
 end if
 if 110 ≤ r ≤ 126 /*Case 3*/
   r ← GetPostPrandialBG ()
   if r > 140
       Print "Possible Diabetes"
   else
       Print "Potential Diabetes
   end else
   end if
end if
```

When pre-prandial value is greater than 126, the post-prandial blood glucose value is requested. If post-prandial value is greater than 140 but less than 200, average blood sugar is calculated and then the results are compared to reference values. If the pre-prandial value is less than 110 and post-prandial value is less than 140, the results are compared to reference values and if the results are less than reference value, it means that person is healthy. Another two situations are the same that if the pre-prandial value is greater than 110 but less than 126, post-prandial value is less than 140, and if the pre-prandial value is greater than 126, post-prandial value is between 140 and 200, the results are compared to reference values, and it means that the person will be potential diabetes and average blood sugar is calculated.

As the third case detailed in Algorithm 2, if the pre-prandial value is greater than 126, the post-prandial blood glucose value is greater than 200, person is a potential diabetes. In order to diagnose the patient, SVM that is soft computing technique is used as classifier. This classifier uses 9 different physiological parameters such as; Pregnancy, glucose, blood pressure, skin fold, insulin, Hemoglobin A1C, body mass index, family tree and age to determine a prediction about the possibility of being diabetic.

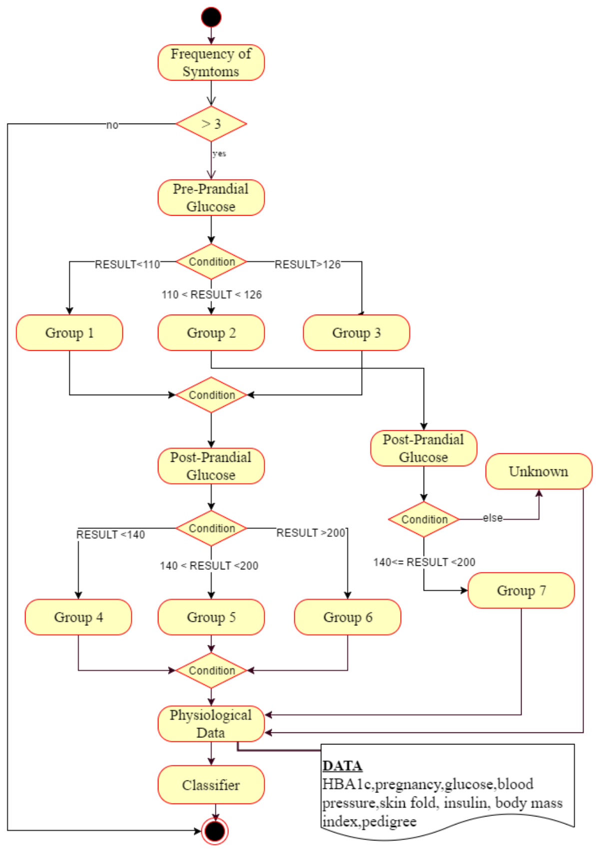


Figure 1. Activity Diagram of Diagnosis Process

4. Experiments and Results

Our reference data is consisting of 242 patients' physiological parameters. We used 10-fold cross validation technique to train and test classifier in our experiments.

The first factor in the success of the selected classifier is the effectiveness of the identified attributes. The goal in attribute selection is to find the subclass that most effectively expresses an achievement measure. Various search approaches such as genetic algorithm are used in the identification of subclass. Among these approaches are mostly preferred; Best Individual Feature, Sequential Forward Search, Sequential Floating Forward Search, Plus-L-Minus-R.

In recognition problems, each feature component has a certain recognition coefficient. The effects of these feature components should be analyzed in order to increase the performance of the recognition. One of the simplest way to learn the attributes' components for recognition is to evaluate the recognition performance of each component separately. For the best recognition, the set of attributes consisting of the highest coefficients should be determined and classified. In the advanced selection search algorithm, the highest coefficient components are included in the attribute set. At each iteration, the contribution of the newly included component to the overall performance of the system is validated to determine whether it is in the cluster.

The experiments are composed at two phases. In the first step, all values of attributes belonging to the database were classified with using SVM and the results were evaluated. In the second phase, the number of inputs are reduced by using the Sequential Forward Search (SFS) algorithm and the classification is repeated with the help of the SVM method.

Experiments using SFS have used 4 different kernel functions such as Linear, Polynomial, Gaussian Radial Basis Function Kernel (GRBFK) and Multilayer Perceptron Kernel (MPK). Physiological parameters were indexed as pregnancy (1), glucose (2), blood pressure (3), skin fold (4), insulin (5), body mass index (6), pedigree (7) and The reduced subsets of the indexed attributes in the specified kernel functions are shown in Table 1.

Table 1 – Reduced attribute sub-sets according to Kernel Functions

Kernel	Attribute Sub-sets
Linear	1,2,4,6,7
Polynomial	1,2,3,6,7
GRFBK	1,2,5, 6
MPK	1,2,3,5,6,7

Accuracy, sensitivity, specificity, ROC domain, and F-measure parameters were chosen as criteria in the analysis of the classification success. The functions used in the calculation of these parameters follow as;

$$\text{Accuracy} = (TP + TN)/\,(TN + TP + FP + FN)$$

$$\text{Sensitivity} = TP/(TP + FN)$$

$$\text{Specificity} = TN/(TN + FP)$$

$$\text{F Measure} = 2 * (\text{Accuracy} * \text{Sensitivity})/(\text{Accuracy} + \text{Sensitivity})$$

Sensitivity and originality are the parameters used to verify the classification ability of the system. Sensitivity is calculated by matching the correct samples to the correct classes. Accuracy is known as the percentage of correct matching. Table 2 presents the classification performance of the two stages based on the criteria. When the results are examined, it is seen that Linear kernel function has higher success rates than other functions before and after the feature reduction.

Table 2 – The Accuracy, Sensitivity, Specificity, ROC Domain and F Measure values of Kernel Functions

	Accuracy		**Sensitivity**		**Specificity**		**ROC Domain**		**F Measure**		**Confusion Matrix**			
Attribute Reduction	Off	On	Off	On	Off	On	Off	On	Off	On	Off		On	
Linear	0,78	0,78	0,81	0,78	0,72	0,77	0,76	0,80	0,76	0,77	TP 180	FN 33	TP 172	FN 27
											FP 41	TN 84	FP 49	TN 90
Polynomial	0,74	0,77	0,75	0,77	0,71	0,76	0,62	0,77	0,73	0,76	TP 166	FN 34	TP 171	FN 28
											FP 55	TN 83	FP 50	TN 89
GRBFK	0,72	0,72	0,71	0,75	0,74	0,68	0,5	0,5	0,72	0,71	TP 158	FN 31	TP 165	FN 37
											FP 63	TN 86	FP 56	TN 80
MPK	0,67	0,70	0,64	0,70	0,71	0,73	0,42	0,51	0,68	0,71	TP 142	FN 33	TP 154	FN 32
											FP 79	TN 84	FP 67	TN 85

The criterion values of the ROC curves (AUC) of SVM classifier are given in Figures 1, 2, 3 and 4 respectively. In the first test SVM is used with Linear Kernel. Results showed that the Attribute Reduction (AR) technique, increases the classification rate by 3% on Figure 2.

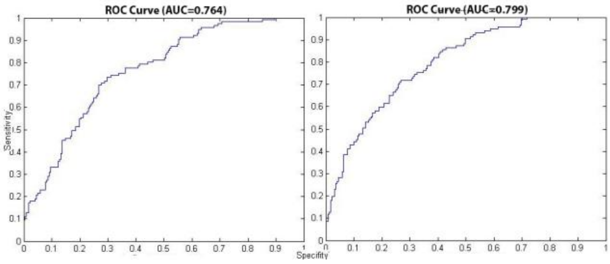


Figure 2. ROC Curves that are computed with Linear Kernel (a) Advanced lookup search algorithm w/o using AR (b) with using AR

Polynomial Kernel exhibits much appreciable success with AUCs of AR-On (0,77) and AR-Off (0,62) are shown on Figure 3.

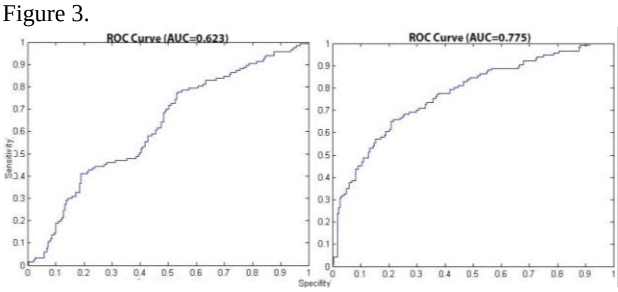


Figure 3. ROC Curves that are computed with Polynomial Kernel (a) Advanced lookup search algorithm w/o using AR (b) with using AR

The usage of with Gaussian Radial Basis Function Kernel plays no significance on classification even with RA.

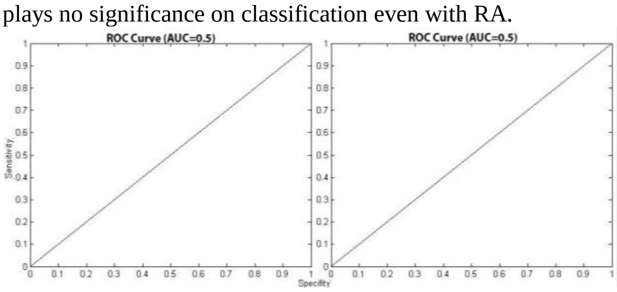


Figure 4. ROC Curves that are computed with Gaussian Radial Basis Function Kernel (a) Advanced lookup search algorithm w/o using AR (b) with using AR

As the last case study, Multilayer Perceptron Kernel is used to demonstrate the classification performance as displayed on Figure 5. AUCs of the AR-On (0,51) results in higher value according to AR-Off (0,42).

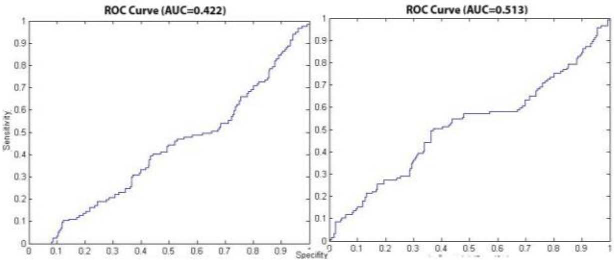


Figure 5. ROC Curves that are computed with Multilayer Perceptron Kernel (a) Advanced lookup search algorithm w/o using AR (b) with using AR

Referring to all four case studies, we showed that enabling AR increase the classification rate in a positive way. The highest dissimilitude is gained with using Polynomial Kernel with AR. However, Linear Kernel with AR gives the best AUS as (0,79).

5. Conclusions

This article has aimed at making a prediction about the diabetes and thanks to Clinical Decision Support System we developed, the diagnosis can be established by getting the patients information and comparing them reference data coming from the hospitals with rapidly.

A model that classifies patients' physiological data to predict the possibility of being diabetic was described. Patient’s physiological parameters such as Hemoglobin A1C, Pre-prandial Blood Glucose, Past-prandial Blood Glucose and Glucose in Urine are requested to compare and decide whether the patient is diabetic or healthy. These inputs guide the system into three level prediction. First two levels generate the output with traditional computing methods. For the patients that are fell into third category, system uses a soft computing approach to compute prediction rates instead of traditional comparisons.

Several tests have been applied to give success rate of the achievement in the third category so that numerous learning algorithms in Support Vector Machine concept were used. In the proposed method, 9 physiologic parameters (Pregnancy, glucose, blood pressure, skin fold, insulin, Hemoglobin A1C, body mass index, family tree and age) of 242 diabetic patients were used. In these experiments in which the support vector machines are run as classifiers, the performance subdivisions of the 9 attributes and the feature subset of the forward search algorithm are classified separately and the performance analysis is performed. It is determined that the system predict the patients' possibility of being diabetic with around 80% success rate.

As a future work, it is aimed to train the system with more verified data to increase prediction rate and optimizations for the computational parts are planned.

6. References

[1] Barakat, N.; Bradley, A.P.; Barakat, M.N.H., "Intelligible Support Vector Machines for Diagnosis of Diabetes Mellitus," Information Technology in Biomedicine, IEEE Transactions on , vol.14, no.4, pp.1114,1120, July 2010

[2] UK Hypoglycaemia Study Group. "Risk of hypoglycaemia in types 1 and 2 diabetes: effects of treatment modalities and their duration." Diabetologia 50.6 (2007): 1140-1147.

[3] American Diabetes Association. "Diagnosis and classification of diabetes mellitus." Diabetes care 31.Supplement 1 (2008): S55-S60.

[4] Huxley, Rachel, Federica Barzi, and Mark Woodward. "Excess risk of fatal coronary heart disease associated with diabetes in men and women: meta-analysis of 37 prospective cohort studies." Bmj 332.7533 (2006): 73-78.

[5] Alavi, Maryam, and John C. Henderson. "An evolutionary strategy for implementing a decision support system." Management Science 27.11 (1981): 1309-1323.

[6] Kim, K. Kyu, and Jeffrey E. Michelman. "An examination of factors for the strategic use of information systems in the healthcare industry." MIS Quarterly (1990): 201-215.

[7] Hunt, Dereck L., et al. "Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review." Jama 280.15 (1998): 1339-1346.

[8] Sim, I., Gorman, P., Greenes, R. A., Haynes, R. B., Kaplan, B., Lehmann, H., & Tang, P. C. (2001). Clinical decision support systems for the practice of evidence-based medicine. Journal of the American Medical Informatics Association, 8(6), 527-534.

[9] Kostic, P.; Vasiljevic, Z.; Pavlovic, S.; Milosavljevic, I.; Milanovic, J.G.; Blesic, S.; Milanovic, S., "Knowledge management system for clinical decision support — Application in cardiology," Telecommunications Forum (TELFOR), 2011 19th, pp.1261,1264, 22-24 Nov. 2011.

[10] Andres El-Fakdi, Francisco Gamero, Joaquim Meléndez, Vincent Auffret, and Pascal Haigron. 2014. eXiTCDSS: A framework for a workflow-based CBR for interventional Clinical Decision Support Systems and its application to TAVI. Expert Syst. Appl. 41, 2 (February 2014), 284-294.

[11] Kamaleswaran, R.; McGregor, C., "Integrating complex business processes for knowledge-driven clinical decision support systems," Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE , vol., no., pp.1306,1309, Aug. 28 2012-Sept. 1 2012.

[12] Reeda Kunhimangalam, Sujith Ovallath, and Paul K. Joseph. 2014. A Clinical Decision Support System with an Integrated EMR for Diagnosis of Peripheral Neuropathy. J. Med. Syst. 38, 4 (April 2014), 1-14.

[13] Shannon Standridge, Robert Faist, John Pestian, Tracy Glauser, and Richard Ittenbach. 2014. The Reliability of an Epilepsy Treatment Clinical Decision Support System. J. Med. Syst. 38, 10 (October 2014), 1-6.

[14] 14. Andrew King, Alex Roederer, Sanjian Chen, Nicholas Stevens, Philip Asare, Oleg Sokolsky, Insup Lee, Margaret Mullen-Fortino, and Soojin Park. 2010. Demo of the Generic Smart Alarm: a framework for the design, analysis, and implementation of smart alarms and other clinical decision support systems. In Wireless Health 2010 (WH '10). ACM, New York, NY, USA, 210-211.

[15] Juha Kemppinen, Jukka Korpela, Roce Partners, Kalle Elfvengren, Timo Salmisaari, Jussi Polkko, and Markku Tuominen. 2013. A Clinical Decision Support System for Adult ADHD Diagnostics Process. In Proceedings of the 2013 46th Hawaii International Conference on System Sciences (HICSS '13). IEEE Computer Society, Washington, DC, USA, 2616-2625.

[16] Shazia Karim and Imran Sarwar Bajwa. 2011. Clinical Decision Support System Based Virtual Telemedicine. In Proceedings of the 2011 Third International Conference on Intelligent Human-Machine Systems and Cybernetics - Volume 01 (IHMSC '11), Vol. 1. IEEE Computer Society, Washington, DC, USA, 16-21.

[17] Eren Gultepe, Hien Nguyen, Timothy Albertson, and Ilias Tagkopoulos. 2012. A Bayesian network for early diagnosis of sepsis patients: a basis for a clinical decision support system. In Proceedings of the 2012 IEEE 2nd International Conference on Computational Advances in Bio and medical Sciences (ICCABS '12). IEEE Computer Society, Washington, DC, USA, 1-5.

[18] Mattila, J.; Koikkalainen, J.; Virkki, A.; van Gils, M.; Lötjönen, J., "Design and Application of a Generic Clinical Decision Support System for Multiscale Data," Biomedical Engineering, IEEE Transactions on , vol.59, no.1, pp.234,240, Jan. 2012.

[19] The Rights of Patient Law (Numbered 23420,21th Clause) http://www.saglik.gov.tr/TR/belge/1-555/hasta-haklari-yonetmeligi.html

[20] International Expert Committee. "International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes." Diabetes care 32.7 (2009): 1327-1334.

[21] El-Nabarawy, I.; Abdelbar, A.M.; Wunsch, D.C., "Levenberg-Marquardt and Conjugate Gradient methods applied to a high-order neural network," Neural Networks (IJCNN), The 2013 International Joint Conference on , pp.1,7, 4-9 Aug. 2013.

[22] Tai-cong Chen; Da-jian Han; Au, F.T.K.; Tham, L.G., "Acceleration of Levenberg-Marquardt training of neural networks with variable decay rate," Neural Networks, 2003. Proceedings of the International Joint Conference on , vol.3, no., pp.1873,1878 vol.3, 20-24 July 2003.

[23] Hannan Ma; Husheng Li, "Analysis of Frequency Dynamics in Power Grid: A Bayesian Structure Learning Approach," Smart Grid, IEEE Transactions on , vol.4, no.1, pp.457,466, March 2013.

[24] Yunong Zhang; Bingguo Mu; Huicheng Zheng, "Link Between and Comparison and Combination of Zhang Neural Network and Quasi-Newton BFGS Method for Time-Varying Quadratic Minimization," Cybernetics, IEEE Transactions on , vol.43, no.2, pp.490,503, April 2013.

[24] Moh, Y.; Buhmann, J.M., "Manifold regularization for semi-supervised sequential learning," Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on , vol., no., pp.1617,1620, 19-24 April 2009

Fatma Patlar Akbulut received her BS degree in 2015 and MS degree in 2009 from Istanbul Kültür University. She started her PhD studies in 2013 in Istanbul University and she's still continuing her studies. As professional experience, she worked as senior software developer, business analyst, and system analyst in Bilgi Sistemleri, Aras Holding, and Ericsson respectively. Her research areas are signal processing, software engineering, pattern recognition and smart wearable systems.

Aydın Akan received his Ph.D. degree in Electrical Engineering from the University of Pittsburgh, Pittsburgh, PA, USA, in 1996. He has been with the Department of Electrical and Electronics Engineering, Istanbul University since 1996 where he currently holds a Professor position. His research interests are non-stationary signal processing, time–frequency signal analysis methods and their applications to wireless communications and biomedical engineering. Dr. Akan is a senior member of the IEEE Signal Processing Society and editorial board member of the Digital Signal Processing journal.

Microcontroller Based Wye-Delta Starter and Protection Relay for Cage Rotor Induction Motor

Mehmet Onur GULBAHCE[1], Aysel ERSOY YILMAZ[2], Derya Ahmet KOCABAS [1]

[1] Department of Electrical Engineering, Faculty of Electrical and Electronics Engineering, Istanbul Technical University, Istanbul, Turkey
[2]Department of Electrical & Electronics Engineering, Faculty of Engineering, Istanbul University, Istanbul, Turkey
ogulbahce@itu.edu.tr, aersoy@istanbul.edu.tr, kocabasde@itu.edu.tr

Abstract: *In this study, PIC 18F4520 based wye-delta starting and protection relay was designed and manufactured with semiconductor devices for three phase cage rotor induction motors. Protection relay controls the phase absence and phase sequence before motor starting process. Designed relay system performs better than that of other electromechanical starting and protection relays, due to the decrease in the number of electromechanical components, total volume of relay, complexity of system. Furthermore, the use of microcontroller and semiconductor devices makes the new starter and protection relay design more robust and more reliable.*
Keywords: *Wye-delta starting, motor protection, PIC 18F4520, cage rotor induction motor*.

1. Introduction

Since a large amount of energy is required to start an electric motor to overcome the inertia of whole drive system, the starting current of a cage rotor induction motor (CRIM) is increased by 2-8 times than the rated current. Starting current at direct on-line start depends on the rated power of motor and the effective rotor resistance at starting conditions [1-3]. Due to high current demand from the network, a sequence of problems such as voltage drop, high transients and, in some cases, even uncontrolled power-cut can occur during starting [1]. A torque impulse that causes considerable mechanical stress on the rotor bars and windings is also generated [3-5]. Besides, in literature and practice, there are numerous different starting methods, namely, auto-transformer start, soft-start, frequency start, rheostat start, star-delta start etc. and all mentioned methods aim to diminish these undesirable effects at start-up [1, 3].

One of the most common and economic starting methods is wye-delta starting that reduces the winding voltage by √3 times and this method is only applicable to delta connected motors. During the starting process, starting current is approximately 30 % and starting torque is approximately 25-30% of direct on line starting. Ordinarily, a wye-delta starter contains at least three contactors with mechanical and electrical interlocks, and a time limit relay which can be seen in Figure 1. 6 terminals of stator windings are connected properly by use of contactors so as to provide wye-connection and then the connection is switched to delta by another contactor while terminating neutral point at the same time. Basically, a time relay is used to adjust the duration of wye connection that effects current and torque impulse which are reduced as the motor gets closer to synchronous speed. The starting time which can also be determined practically is affected by load torque, starting torque and the inertia of drive system.

Figure 1. Wye-delta starter with mechanical and electrical interlocks, and a timing system.

Since electrical and mechanical interlocks of contactors abrade over time, mechanical failure, increase in maintenance cost and noise are inevitable over time. In addition, the large numbers of electromechanical components, high volume and the complexity of the relay system reduce starting performance, robustness and reliability [6, 7].

Received on: 13.01.2017
Accepted on: 09.03.2017

In this study, a more reliable, robust, maintenance-free, noiseless and high-performance wye-delta starter is designed and manufactured by means of a microcontroller and semiconductor devices. Moreover, designed starter circuit also protects CRIM from phase absence and fault of phase sequence with the help of a protection relay algorithm which is embedded into the microcontroller.

2. Materials and Methods

Microcontroller based wye-delta starter and protection relay consists of three sub-systems which are wye-delta starter, phase protection relay for CRIM and microcontroller together with its interfaces.

2.1. Wye-Delta Starter for CRIM

The wye-delta starters are very common winding voltage reducing starters in industrial applications. The method aims to decrease the starting current by reducing the applied winding voltage. This also reduces the disturbances to the network. In many networks, direct on-line starting for the motors with a rated power greater than 4 kW is restricted by regulations [8]. Wye-delta starter is one of the lowest cost voltage reducing starters which is applicable only to delta connected motors in rated operation.

The electromechanical wye-delta starter contains at least three contactors, namely main, wye and delta contactors. The system contains also mechanical and electrical interlocks and a timing system. Figure 2a depicts a basic configuration for a wye-delta starting system with electromechanical equipment. During starting, initially, the main contactor (M) and star contactor (Y) are closed. When the motor gets close to rated speed, star contactor (Y) is opened and delta contactor (Δ) is closed. Control of contactors is provided by mechanical & electrical interlock together with the timing system built into the starter.

One of the main drawbacks of the method is current and torque impulses occurring at two different instances. First disturbance is at the very beginning of starting and the second one is at the instance of connection changing from wye to delta. First impulse happens depending on the nature of starting and second impulse is created at the time that the contacts are opened and closed. The use of electromechanical wye-delta starter creates “Open Transition” during connection change.

Since CRIM is disconnected from the line temporarily while the contacts change position, the connection is being altered from wye to delta occurs. Depending on the transition time, speed of change and the load driven by CRIM, fluctuations in motor current and torque are generated which cause unwanted mechanical and electrical facts on system. In some cases, instantaneous transient current exceeds even the locked rotor current for a short duration and this instantaneous fluctuation can be powerful enough to damage system components. In order to reduce the current and torque fluctuations, the transition time must be reduced. Since electronic switches are faster than electromechanical ones, it is possible to use semiconductor devices instead. A wye-delta starter performed with semiconductor devices such as triac is given in Figure 2b.

A triac is a solid state power switching device that allow current both directions. In other words, a triac can be triggered into conduction by both negative and positive waves and together with both negative and positive trigger pulses applied to its gate terminal. Triacs provide faster, more economical and impeccable power control than mechanical and electromagnetic switches under alternating current. Because of high commutation capability, it does not generate arc during switching operations resulting in mechanical deformations. Cooling requirements are very small due to low on-state voltage drop and dynamic resistance.

In this study, BTA41-600B is used as switching power device which can be frequently used in static relays, heating regulation, induction motor starter, light dimmers and motor speed controllers etc.

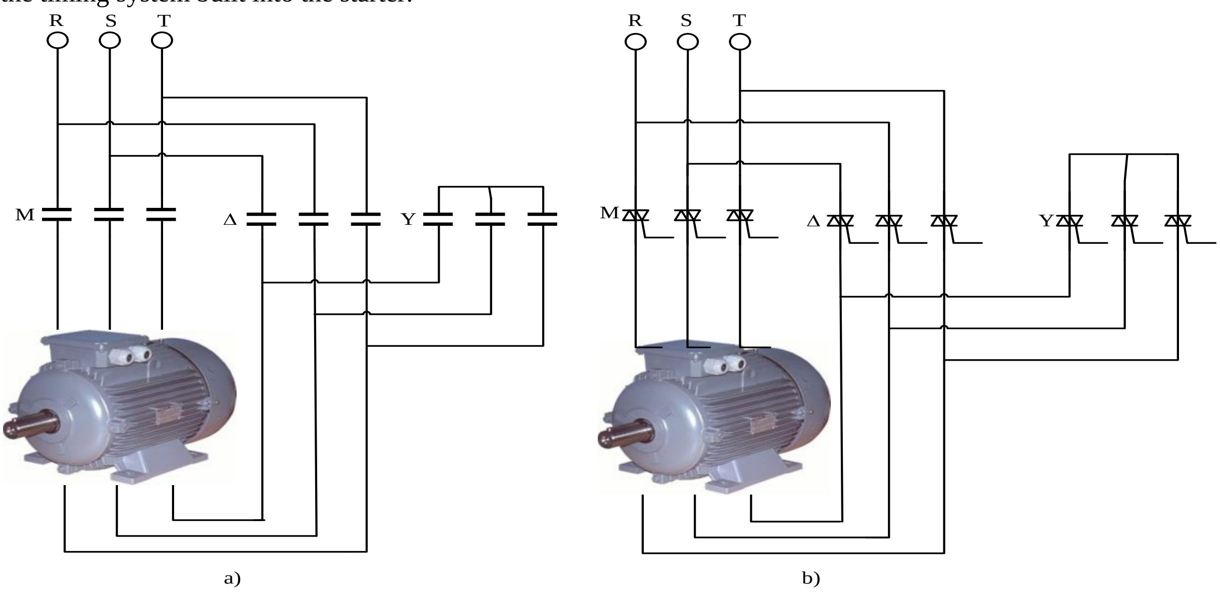


Figure 2. Wye-delta starter connection diagrams a) with electromechanical devices b) with semiconductor devices.

2.2. Microcontroller Based Phase Protection Relay

Microcontroller based phase protection relay provide protection against phase absence and phase sequence (or reverse sequence). Phase absence occurs when one phase or more is off which can be typically caused by a blown fuse, broken wire, loose connection or worn contact. This malfunction leads to unnecessarily high currents and the CRIM can still continue to run even after losing one phase resulting in potential motor burn-out. Reversing any two of phases or fault of phase sequence forces a CRIM to run in opposite direction. This may cause damage, accidents, injury, etc. in some practical cases.

In industrial applications, phase protection relay cannot be connected directly to CRIM. An additional contactor system with mechanical and electrical interlocks must be used. Since contactors are electromechanical devices, the life time and number of opening and closing are definite.

In this study, triacs and a microcontroller are used for phase protection making the system free from electromechanical devices. Absence and sequence (or reverse sequence) of phases connected CRIM terminals were detected by microcontroller. An LCD screen is used to visualize the situation for the user. Operational situation and error messages for the faults are printed on the LCD. Basic principle diagram of protection relay given in Figure 3.

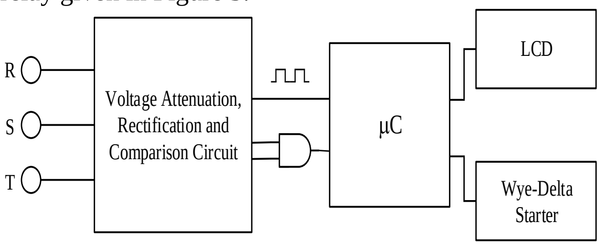


Figure 3. Basic principle diagram of microcontroller (µC) based protection relay.

Voltage signals sensed from three-phase terminals are converted into transformed signals via the "voltage attenuation, rectification and comparison circuit" that can be detected by the microcontroller. It can be seen in Figure 4 that each phase (R, S and T in order) has positive cycles just after the end of the positive cycle of the phase in sequence. Reversal of the phase sequence and the phase fault can be determined by this approximation.

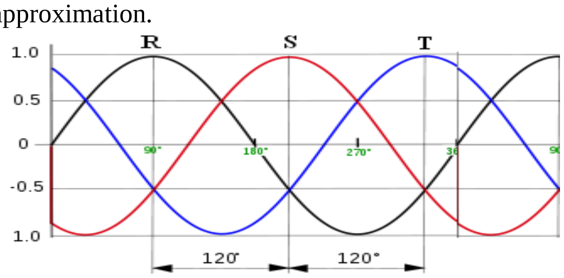


Figure 4. Waveform of three-phase voltages.

In order to obtain more accurate results, three open loop operated operational amplifier (OPAMP) and passive circuit elements which are illustrated in Figure 5 were used. These circuits attenuate the phase voltages and convert into square waves with 120º phase difference. In the figure given below, RE0, RE1 and RE2 are connected to E port of microcontroller to detect malfunction in phase sequence.

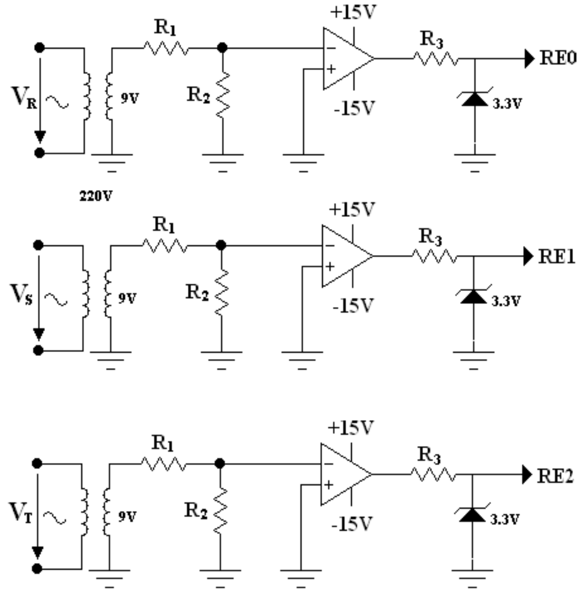


Figure 5. Voltage attenuation and comparison circuits performed by open-loop opamps.

In order to detect phase absence, three-phase network voltages were rectified and attenuated. Rectified and attenuated voltage signals are applied to an AND gate. If all three phases exist, the output of the AND gate will be logic 1. If there is a fault at least at one phase, the output of the gate will be logic 0. Phase absence detection circuit is given in Figure 6. In given diagram output of AND gate (RD7) is connected to microcontroller.

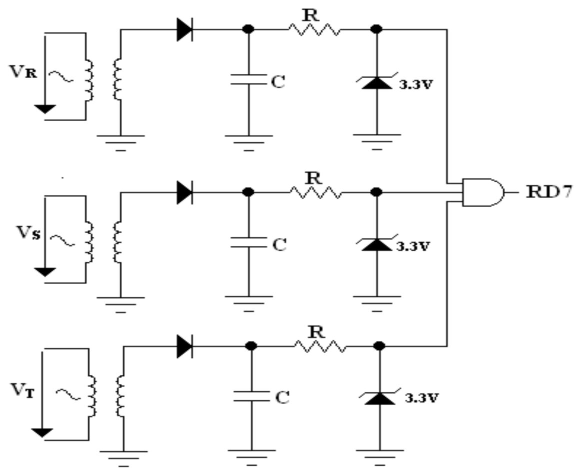


Figure 6. Circuit diagram of voltage attenuation and rectification.

2.3. Microcontroller and Its Interfaces

In this study, PIC 18F4520 microcontroller from the enhanced flash microcontroller family with 10-Bit A/D and

Nano-Watt technology is used. PIC 18F4520 has 40 MHz operating frequency, 32768 Bytes program memory, 1536 Bytes data memory, 256 Bytes EEPROM (Electrically Erasable Programmable Memory), three programmable external interrupts, four input change interrupts, five input and output units (A, B, C, D, E), four timers, one “capture / compare / PWM module and one enhanced capture / compare / PWM module”, serial communication modules, 10-bit analog digital converter module, programmable low / high voltage detection etc [9].

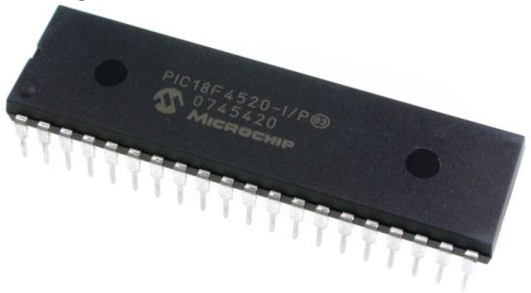


Figure 7. PIC 18F4520 from the enhanced flash microcontroller family with 10-Bit A/D and Nano-Watt technology [9].

Output control signals of microcontroller cannot be used directly to trigger a triac. A triac drive circuit is used both to bring the control signals to a level that can turn-on triacs and to provide an electrical isolation between the microcontroller and starter triac circuit. A simple phase drive circuit with zero voltage crossing bilateral triac driver MOC 3041 which is used in the study is given in Figure 8.

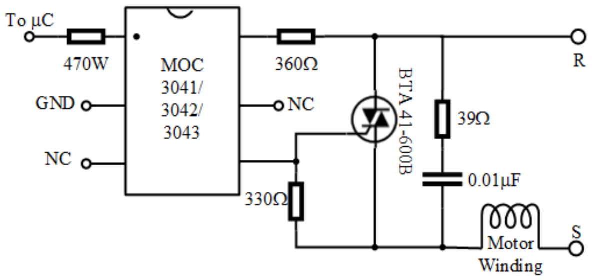


Figure 8. A simple phase drive circuit with zero voltage crossing bilateral triac driver [10].

Microcontroller needs a software to be prepared to process the input signals. Basic flowchart of wye-delta starter and protection relay for CRIM is depicted in Figure 9. Operational situation and faults are displayed on the LCD screen by messages prepared in Turkish.

In the prepared and embedded software, phase absence is checked initially. If there is a fault in phase(s), error message is created on LCD for the operator. Secondly, the phase absence and phase sequence are checked. If all phases exist and phase sequence is true, the CRIM is started in wye connection. The wye-delta switching time is set with a delay function in the prepared software and the time depends on the total inertia of the drive system and practical observation on speeding up. Finally, when the motor reaches around rated speed, motor windings are then connected delta by means of triacs.

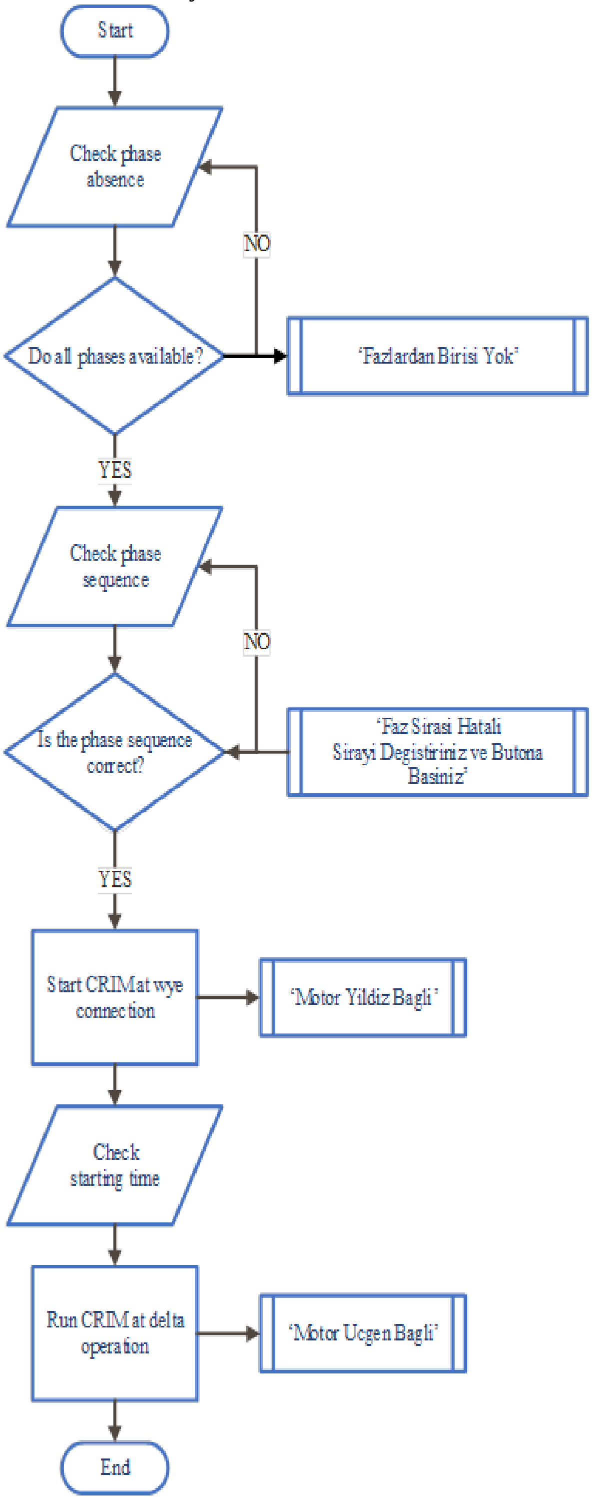


Figure 9. Basic flowchart of wye-delta starter and protection relay for CRIM.

3. Practical Implementation

All sensing and processing sub-circuits are placed on to same circuit board together with other motor protection elements and circuit breakers. All parts are installed to a box and LCD screen is mounted on the box cover. Designed and manufactured microcontroller based wye-delta starter and protection relay for CRIM Figure 10-11.

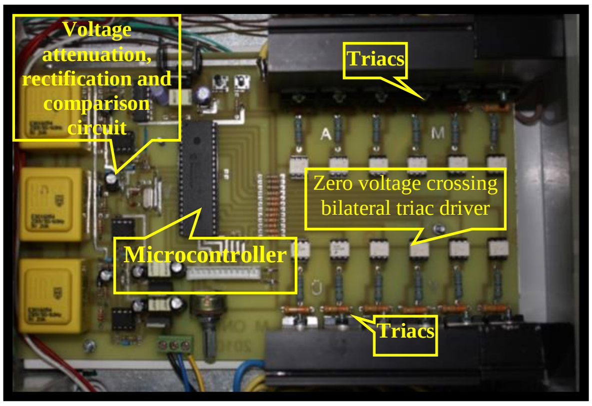


Figure 10. The printed circuit board of designed starter and protection relay and its layout of the elements.

Figure 11. The boxed designed starter and protection relay and its layout of the elements.

4. Test Setup and Application Studies

In order to test to microcontroller based wye-delta starter and protection relay for CRIM, the "FH2 Mk1V" test setup of "TQ Education and Training System" is used with a "FH90 induction motor kit". Motor is loaded by an "Eddy Current Brake" mounted on the bench. The test bench is a modular one that is designed for different types of low power electrical machines. Physical test setup is given in Figure 12.

Necessary connections between CRIM and designed starter were established. The created scenario is to start the motor in wye connection, to run the motor 5 secs and to switch to delta connection by microcontroller based wye-delta starter. When everything is permissible "phase sequence is ok", "motor is wye connected" and "motor is delta connected" (all in Turkish) is written on LCD screen respectively, as seen in Figure 13-14.

During tests, one of the phases was deactivated on purpose, and the starter immediately interrupted CRIM while "one of the phases is missing" (in Turkish) is displayed. (Figure 15) After the phase fault was cleared, the phase sequence is changed for test. "Phase sequence is incorrect, change order and press button" (in Turkish) message is displayed and the motor is not allowed to start. (Figure 16)

CRIM can only be allowed to start in wye connection after clearing all faults.

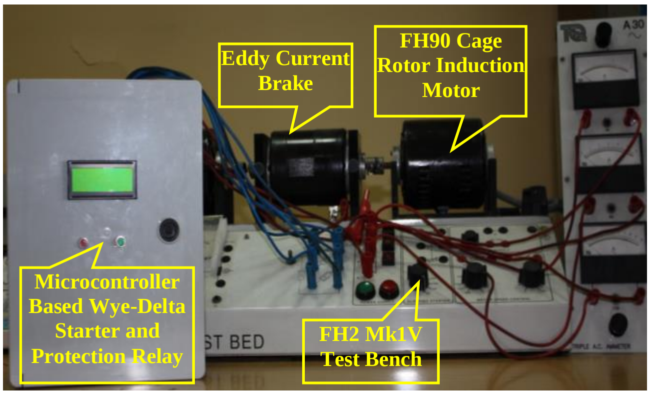


Figure 12. Test setup for designed starter and protection relay.

Figure 13. The state that the phase sequence is confirmed by the microcontroller and CRIM windings are wye connected.

Figure 14. The state that the phase sequence is confirmed by the microcontroller and CRIM windings are delta connected.

Figure 15. The state that the one of the phases is in fault.

Figure 16. The state that the one of the phases is in fault.

5. Conclusion

In this study, microcontroller based wye-delta starter and protection relay for cage rotor induction motors is designed, manufactured and physically tested. The starter and the relay are separated into three main sub-systems and all sub-systems were designed individually and the system is integrated. The sub-systems are “electronic wye-delta starter consisting of triacs”, “microcontroller based phase protection relay” and “microcontroller and its interfaces”.

Nine triacs were used to connect the stator windings of a CRIM instead of three three-phase contactors. Triacs were driven by a phase drive circuit with zero voltage crossing bilateral triac driver.

Sensor circuits for phase sequence and phase absence detection are designed. Phase voltages are attenuated and compared with reference voltage to adapt the converted signals to microcontroller inputs. For evaluating phase absence, measured phase voltages are rectified and applied to an AND gate. All obtained signals are processed by a routine in the microcontroller for detection and visualization. Fault and operational messages are displayed on LCD screen in Turkish.

Practical electronic relay and starter was produced and laboratory tests were performed. Regular start-up procedure without faults was put into practice and expected messages were seen on screen while the motor started according to the prepared scenario. Missing phase and incorrect phase sequences were created and it was seen that the motor was protected by the electronic relay and all expected error messages were seen on screen. It was tested that motor was allowed to start after all faults were cleared.

A low-volume, noise-free, robust, reliable, long-life electronic starter and a protection relay was produced. All disadvantages of mechanical and electromechanical starters and relays were overcome. A faster system response was provided by replacing electromechanical switches and a faster transition interval is provided, since mainly mechanical time-constant is highly greater than electrical one. Acoustic noise was eliminated by using semi-conductor switches. A maintenance-free and a robust starter was provided by removing mechanical contacts. Also the life of the starter was extended by implementing soft-switches. Total volume of the starter was reduced. Overall, system reliability was increased.

5. References

[1] Chapman Stephen J., *Electric Machinery Fundamentals*, McGraw-Hill, New York, 1991.

[2] Fitzgerald A. E., Charles Kingsley, and Stephen D. Umans., *Electric Machinery*. McGraw-Hill, Boston, 2003.

[3] Mergen A. F., Zorlu S., *Electrical Machines-II-Asynchronou Machines*, Birsen Publisher, Istanbul, 2009 (In Turkish)

[4] Mergen A. F., Kocabas D. A., *Windings in Electrical Machines*, Birsen Publisher, Istanbul, 2009 (In Turkish)

[5] Gieras J. F., *Electrical Machines: Fundamentals of Electromechanical Energy Conversion*, CRC Press, Portland, 2016.

[6] Colak I., Bayindir R., “*Implementation of the Classical Starting Techniques Using Microcontroller”*, Journal of Science & Engineering of Firat University, Vol: 17, No: 1, pp. 10-18, 2005.

[7] Bayindir R., Demirbas S., Irmak E., Bekiroglu E., *“Design and Implementation of Microcontroller Based Starting and Protection Relay for Induction Motors”*, Journal of Polytechnic, Vol: 10, No: 1, pp. 1-5, 2007.

[8] Regulations of Internal Electrical Installation, Retrieved January 12, 2017, from *http://www.emo.org.tr/ekler/cc5a85842d8d326_ek.pdf.*

[9] PIC18F2420/2520/4420/4520 Data Sheet, Retrieved January 12, 2017, from *http://ww1.microchip.com/downloads/en/devicedoc/39631a.pdf*

[10] MOC 3041 Data Sheet, Retrieved January 12, 2017, from *https://www.fairchildsemi.com/datasheets/MO/MOC3041M.pdf*.

Mehmet Onur Gulbahce received the B.S. degree in electrical engineering from Istanbul University, Istanbul, Turkey, in 2010. He received the M.Sc. degree in 2013 in Electrical Engineering Programme, Institute of Science and Technology, ITU and since 2013 he has been a Ph.D. student in the same programme. He is also a research assistant in the Department of Electrical Engineering, Electrics and Electronics Faculty, ITU since 2011. His main subjects of concern are high-speed electrical machines, drive system and dc-dc converters.

Aysel ERSOY YILMAZ received her M.Sc. and Ph. D. Degrees in Electrical and Electronics Engineering from Istanbul University in 2003 and 2007, respectively. She has published more than 30 national an international papers in journals and conferences. Since spring 2008, she has been working as as Assistant Professor at the Electrical and Electronics Engineering Department of Istanbul University.

Derya Ahmet Kocabas received the B.S. degree in electrical engineering from ITU, Istanbul, Turkey, in 1994. He received the M.Sc. and Ph.D. degrees from Electrical Engineering Programme, Institute of Science and Technology, ITU, in 1997 and 2004, respectively. His main subjects of concern are design and control of electrical machines, space harmonics, drive systems and power electronics. He joined to Department of Electrical Engineering, Electrics and Electronics Faculty, ITU in 1995 and since January 2009 he has been an Assistant Professor.

HEART SOUND LOCALIZATION AND REDUCTION IN TRACHEAL SOUNDS BY GABOR TIME-FREQUENCY MASKING

Esra SAATÇI[1], Aydın AKAN [2]

[1]Electrical&Electronics Engineering Department, Istanbul Kultur University, Turkey
[2] Department of Biomedical Engineering, Izmir Katip Celebi University, Cigli, Izmir, Turkey
esra.saatci@iku.edu.tr, aydin.akan@ikc.edu.tr

Abstract: *Respiratory sounds, i.e. tracheal and lung sounds, have been of great interest due to their diagnostic values as well as the potential of their use in the estimation of the respiratory dynamics (mainly airflow). Thus the aim of the study is to present a new method to filter the heart sound interference from the tracheal sounds. Tracheal sounds and airflow signals were collected by using an accelerometer from 10 healthy subjects. Tracheal sounds were then pre-processed by Recursive Least Square - Adaptive Noise Cancellation (RLS-ANC) filter to remove background noise. Gabor time-frequency expansion was used for both heart sound localization and reduction problem. In the first step of filtering, RLS-ANC successfully filtered out the broad - band ambient noise. Reconstruction of tracheal sound was achieved from modified Gabor coefficients without heart sound noise. Visual inspection and quantitative analysis demonstrated that Gabor time-frequency masking with RLS-ANC filters provides successful tracheal sound signal separation.*
Keywords: *tracheal sound filtering, Gabor expansion, time-frequency filtering.*

1. Introduction

Respiratory sounds, i.e. tracheal and lung sounds have been of great interest due to their diagnostic values as well as their potential to be used in the estimation of the respiratory dynamics (mainly airflow) [1,2]. The first step to utilize the respiratory sounds is to remove any noise contaminating the valuable spectra-temporal bands of the signals, such that high energy heart sounds interfere with the low energy respiratory sounds at the low frequency band [2]. Heart sound interference should be removed from the respiratory sound signal completely and also efficiently, i.e. without losing or harming the respiratory sound signal overlapping with the heart sound effected frequency band. This is required due to two important factors:

i. Low energy band of the respiratory sound is proved to contain valuable diagnostic information [1],

ii. In order to benefit from respiratory sounds to estimate various respiratory parameters, mainly respiratory airflow and breathing frequency, it is required to work on the clean respiratory sound signal with distinguishable inspiration and expiration parts [2-4].

Statistical signal processing methods were proposed to filter out any noise in the respiratory sound signal, as well as, different approaches based on adaptive filtering [5-7] and time-frequency filtering [7,8] were applied to remove the heart sound noise without altering the respiratory sound signal. All these promising methods achieved a significant degree of success. However, although heart sounds were localized perfectly with these methods, filtering of the respiratory sound signal was still on debate. Moreover, implementations of these methods were not straight forward and needs careful computerization, thus not suitable for automated diagnosis systems.

The need for a fast and effective method can be overcome by the Gabor type time-frequency representation of the respiratory sounds. Gabor representation of a signal provides a convenient means to modify the signal in the time-frequency domain. By adjusting the magnitude of the Gabor coefficients in a prescribed manner and reconstruction of the modified signal using the inverse Gabor expansion, time-frequency filter is easily implemented. In our previous study [9], time-frequency masking technique based on Gabor expansion was applied successfully for the respiratory sound noise reduction problem. However, in [9] tracheal sound signal was used from database as a respiratory sound signal. Given the fact that tracheal sounds from database are pre-processed and cannot represent the real life situation, raw recorded data should be processed. However, raw respiratory signals include not only heart sound signals but also broad band ambient noise. Therefore, the goal of this paper is to evaluate the use of Recursive Least Square - Adaptive Noise Cancellation (RLS-ANC) filter to remove background noise and to assess the effectiveness of Gabor time-frequency masking techniques for heart sound noise localization and reduction problem. Transient noises such as speech and impulsive noise are out of scope of this study.

Received on: 16.01.2016
Accepted on: 21.03.2017

2. Materials and Methods

2.1. Discrete Gabor Expansion

The Gabor expansion and Gabor transform is the time domain - to - time-frequency domain linear and two sided transformation of the signals. By applying sampling theory to continues-time Gabor expansion, discrete Gabor expansion of a finite (or periodic) discrete-time sequence with length L can be defined as [10]:

$$f\left(k\right)=\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}c_{m,n}g_{m,n}\left(k\right),\ k=0,\cdots L-1 \quad (1)$$

where synthesis function $g_{m,n}(k)$ is the $m\Delta M$ time shifted and $n\Delta N$ frequency modulated version of the Gabor window function $g(k)$ also called logons or atoms:

$$g_{m,n}\left(k\right)=g\left(k-m\Delta M\right)e^{\frac{j2\pi n\Delta Nk}{L}} \quad (2)$$

In Eq (2) ΔM and ΔN are time and frequency sampling intervals, respectively. M and N are the number of sampling points in time and frequency domains ($\Delta MM = \Delta NN = L$). Ideally, $g(k)$ should be well localized in both time and frequency (i.e., should decay rapidly outside a small region in the time-frequency space), in which case the coefficients $c_{m,n}$ give good indications of the content of the signal at time $m\Delta M$ and frequency $n\Delta N$. Originally, the synthesis function was chosen by Gabor as a Gaussian window, because it maximizes the concentration in the time-frequency plane

The existence of (1) has been found to be possible for arbitrary $f(k)$ only for $\Delta M\Delta N \le L$ (or $MN \ge L$). This is called the oversampled case and the synthesis functions are no longer linearly independent. At the critical sampling case $\Delta MM = \Delta NN = L$, the logons are linearly independent, but are not orthogonal in general (Balian-Low obstruction) [10,11]. This means that the Gabor coefficients $c_{m,n}$ are not simply the projections of $f(k)$ onto the corresponding logons $g(k)$ (i.e. the analysis and synthesis windows cannot be the same).

According to [12] Gabor coefficient $c_{m,n}$ is computed by the inner product rule for projecting $f(k)$ onto $\gamma(k)$, auxiliary function, or bi-orthogonal window, i.e.,

$$c_{m,n}=\sum_{k=0}^{L-1}f\left(k\right)\gamma^{*}_{m,n}\left(k\right) \quad (3)$$

where again analysis function $\gamma_{m,n}(k)$ is the $m\Delta M$ time shifted and $n\Delta N$ frequency modulated version of the Gabor window function $\gamma(k)$:

$$\gamma_{m,n}\left(k\right)=\gamma\left(k-m\Delta M\right)e^{\frac{j2\pi n\Delta Nk}{L}} \quad (4)$$

Analysis function $\gamma(k)$, is also called dual window function of the synthesis window function since they can be interchangeable. If the windows $g(k)$ and $\gamma(k)$ are chosen biorthonormal, transform is called an orthogonal-like Gabor transform [10] and they validate the biorthonormal condition

$$\sum_{k=0}^{L-1}g\left(k+mN\right)e^{-\frac{j2\pi nMk}{L}}\gamma^{*}\left(k\right)=\delta\left(m\right)\delta\left(n\right) \quad (5)$$

where $0 \le m \le \Delta N - 1$ and $0 \le n \le \Delta M - 1$. Then the analysis formula given by Eq (3) allows the computation of the Gabor coefficients and the synthesis formula in Eq (1) the reconstruction of the signal $f(k)$. Zak transform can be used to compute the biorthonormal window $\gamma(k)$ associated to a given synthesis window $g(k)$. From an implementation point of view, this solution is not fully satisfactory since the computation of the biorthonormal window $\gamma(k)$ is numerically unstable. So in general, some degree of oversampling is considered, which introduces redundancy in the coefficients, in order to "smooth" the biorthonormal window $\gamma(k)$, for the sake of numerical stability. These considerations are closely connected to the theory of frames [13].

2.1.1 Computation of discrete orthogonal-like Gabor expansion

Eq.(1) and Eq. (3) can be expressed in the matrix form respectively as

$$\mathbf{f}=\mathbf{Gc} \quad (6)$$

$$\mathbf{c}=\mathbf{fW}^{*} \quad (7)$$

where the sequence **f** is expressed in the form of a column vector:

$$\mathbf{f}=\left[\,f_0\ \ f_1\cdots\ f_{L-1}\right]^{T} \quad (8)$$

G denotes the $L \times MN$ Gabor synthesis matrix having $g_{m,n}$ as its $(m + nM)$-th column, such that

$$\mathbf{G}=\begin{bmatrix} g_{0,0}\left(0\right) & \cdots & g_{M-1,N-1}\left(0\right) \\ \vdots & \ddots & \vdots \\ g_{0,0}\left(L-1\right) & \cdots & g_{M-1,N-1}\left(L-1\right) \end{bmatrix} \quad (9)$$

The Gabor expansion coefficients $c_{m,n}$ are written in the form of a column vector **c** of length MN:

$$\mathbf{c}=\begin{bmatrix} c_{0,0} & \cdots & c_{M-1,N-1} \end{bmatrix}^{T} \quad (10)$$

$\mathbf{W}^{*}$ is the complex conjugate of $L \times MN$ analysis matrix constructed as same as **G**

Eq.(5) can also be expressed matrix-vector notation as:

$$\mathbf{HW}^{*}=\mathbf{I} \quad (11)$$

where $\mathbf{I}$ is identity matrix and $\mathbf{H}$ is a $MN \times L$ matrix constructed by [14]:

$$H\left(m\Delta M+n,k\right)=g\left(k+mN\right)e^{-\frac{j2\pi nMk}{L}} \qquad (12)$$

where $0 \leq m \leq \Delta N - 1$ and $0 \leq n \leq \Delta M - 1$.

As it is explained earlier, in the oversampling case, linear system given Eq.(5) is underdetermined and solution that making the shape of $\gamma(k)$ and $g(k)$ as close as possible in the least square sense can be found [10]:

$$\Gamma=\min\sum_{k=0}^{L-1}\left(\frac{g\left(k\right)}{\left\|g\left(k\right)\right\|}-\frac{\gamma\left(k\right)}{\left\|\gamma\left(k\right)\right\|}\right)^{2} \qquad (13)$$

This solution is then equated to the solution of the system Eq.(10) via pseudo-inverse method, i.e., the window satisfying Eq.(10) with minimal norm is given by:

$$\mathbf{W}^{*}=\mathbf{H}^{*}\left(\mathbf{H}\mathbf{H}^{*}\right)^{-1}\mu \qquad (14)$$

Eq.(14) says that regarding to the oversampling case, biorthogal analysis window function can be easily obtained once the synthesis function is set. Other conclusion can be drawn as the similarity between the pair of dual functions $\gamma(k)$ and $g(k)$ is directly proportional to the oversampling rate.

2.1.2 Denoising by Gabor expansion

Gabor expansion can be used as a tool for a noise reduction, if either the noise components of the signal is well localized and occupies certain number of cells in time-frequency plane [15] or can be assumed that an independently identically distributed Gaussian noise [16]. Acquired respiratory sound signal is composed of many types of noise signals that needed to be filtered off. Heart sound signal can be considered as a periodic type noise signal since its location can be detected easily by any of linear time-frequency signal representations. However, a noise from an electronic measurement circuitry is usually Gaussian type white noise.

Regardless of the type of the noise included in the signal, Gabor coefficients thresholding or modification can be methods used for the noise reduction. Depending on the noise level and statistical properties of the noise, different algorithms are constructed for different tresholding levels [15-17]. In [16] the denoising algorithm was presented in the case of Gaussian type of the noise signal, whereas in [15] time-frequency domain denoising methods were utilized.

Gabor coefficients masking as denoising approach has a fairly simple algorithm. However, the care, that the analysis and synthesis window should be as close as possible, should be taken. Once the constraint of Eq.(11) is satisfied, it is easy to show that the modified Gabor coefficients are closer to the Gabor coefficient of the modified signal via transform Eq.(7) (Proof is in [15]).

2.2. Gabor Representation of the Tracheal Sound

2.2.1 Data Acquisition and preprocessing

In this work respiratory sounds were acquired from 10 healthy subjects in ranging age of 20 to 30 year-old. Respiratory sounds were recorded by 2 accelerometers (PCB 353B16) placed over suprasternal notch and 3rd intercostal space posteriorly on the left. Respiratory air flow was measured by a pneumotachograph (Hans Rudolph RSS 100 0-160 L/min) attached on a facemask (Respironics Spectrum medium size). Subjects were instructed to breathe quietly without making extra effort. The low-noise operational amplifier was used with the gain factor of 5000 for amplification of the raw sound signals. Preprocessing also included RC band pass filter with the bandwidth of 7.5 Hz. to 2500 Hz. The signals were then digitized by data acquisition board (NI PCI-6221 M 16-bits). The sampling rate was 10 kHz. Acquired signals were displayed and saved for processing by data acquisition software (NI LabVIEW full development system).

2.2.2 Ambient noise filtering by RLS-ANC Adaptive Filter

Acquired data did not only contain heart sound and respiratory sound signals but also were affected by the ambient noise and the noise from electronic components. It has been proved that ordinary band pass filters were not useful in terms of noise reduction in the respiratory sound [1, 5, 8]. Thus Recursive Least Squares Adaptive Noise Cancellation (RLC-ANC) was used to filter out the ambient noise in the respiratory sounds.

The standard RLS adaptive filtering scheme consists of a finite-duration impulse response transversal filter and RLS algorithm, which upgrades the tap weights w_k of the transversal filter in a recursive manner so that the cost function is minimized [18]. The details of the RLS-ANC algorithm can be found in [9]. As a reference noise signal, the unconnected accelerometer output was recorded. As the RLS adaptive filter is highly sensitive to numerical instability [18], the filter order severely affects the performance of the filter. In order to keep computational time as low as possible, RLS-ANC filter order was chosen to be 8 on trial and error basis and the changes occurred at the spectrogram of the signals were observed. λ was set to 1 to be infinite memory.

2.2.3 Gabor analysis of respiratory sounds

After adaptive filtering, acquired tracheal sound signal includes both desired tracheal sound signal and heart sound noise. In this work, we used the generalized Gabor expansion for the respiratory sound signal modeling. For finite discrete-time signals, Gabor synthesis and analysis equations are given in Eq.(6) and Eq.(7).

For the synthesis function $g(k)$, we chose Gaussian type window in order to obtain well localized windows [11]. Below normalized Gaussian function is used:

$$g(k)=\left(\frac{\sqrt{2}}{\sigma}\right)^{1/2} e^{-\pi\left(\frac{k}{\sigma}\right)^2}, \quad 0\le k<L \tag{15}$$

where L is the number of the data samples and σ is the standard deviation of the Gauss function

Selecting the standard deviation highly depends on the wavelength and the line of sight of the signal to be detected. By increasing the wideness, the Gabor expansion emphasizes lower frequencies, whereas higher frequencies of the sound signals can be detected with narrow Gabor expansion windows. Therefore in order to detect low frequency content of the respiratory sound signal we tried relatively high values of σ and 256 point of the window length was found optimum on the heart sound signal frequency band.

Biorthogonal sequence was computed using algorithms explained in section 2.1. Both the window and the biothogonal sequence are illustrated in Figure 1. $L = 10000$ sample segment was used to compute Gabor coefficients and we considered oversampling case with $M = 1000$ and $N = 10000$. The over sampling rate can be calculated as $r = \frac{M\times N}{L} = 1000$. Therefore, due to such a high oversampling rate and pursuing orthogonal-like Gabor transform, both of the window functions are Gaussian type with different amplitudes. Once biothogonal $\gamma(k)$ is determined by Eq.(14) and Eq (15), it is trivial to compute $c_{m,n}$ by Eq.(7).

4. Results

Figure 2 shows the spectrogram of the typical recorded tracheal sound signal from one of the representative subject before ambient noise filtering and the same signal after RLS-ANC adaptive filter. Power spectral density (PSD) plots of the signals are shown in Figure 3. Since, our criterion of successful filtering was to have less coloured spectrogram, as it is shown in Figure 2, the spectrogram of the RLS-ANC filtered sound signal has less noise artefacts. Moreover, heart sound noise components are clearer in the RLS-ANC filtered signal than in the original sound signal. Thus filtering off the background noise attenuates the broad-band noise component of the sound signals while

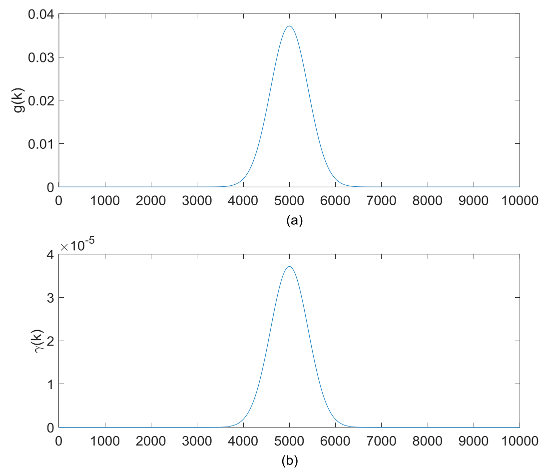


Figure 1. (a) Normalized Gaussian window, $g(k)$ and (b) optimum biorthogonal window, $\gamma(k)$.

exposing more readily sound spectral content. Figure 3 shows the decrease of the PSD in the whole frequency band of the tracheal sounds after the filtering.

Broad-band noise elimination with adaptive filtering was very successful because the ambient noise was uncorrelated with sound signals and cannot be locally identified. However, this is not the case for heart sound signal. Heart sound signal can only be eliminated by the time-frequency representation of the respiratory sound signal. Thus our second approach to the denoising problem was to express the sound signals with linear time-frequency transform.

Figure 4 shows the magnitude values of the Gabor coefficients, $c_{m,n}$ of typical tracheal sound signal in a contour plot before and after denoising. Note that only the positive half of the frequency axis is shown. In Figure 4a, we see that Gabor coefficients are visible only at the frequencies where high energy heart sound signals are present. In other words, most of the Gabor coefficients are close to zero outside the noisy region in the joint time-frequency domain of the tracheal sound signal. This can be explained with two important facts. First, with the selection of the Gaussian window length the low frequency band of the respiratory sound signal is emphasized, and second higher intensities of heart sound made the Gabor coefficient matrix sparse. In other words, coefficients related to heart sound component is too high, so that respiration related coefficients are regarded as zero. This is based on the calculations of the Gabor coefficients, as discussed in section 2.1. Thus, the desired signal can be obtained from the noisy signal by masking the high amplitude Gabor coefficients.

Figure 4b shows Gabor coefficients of the same signal segment after Gabor coefficient masking. As explained in section 2.1.2, applied masking technique is called soft clipping and used when the Gabor coefficients are sparse [16]. Heart sound reduction can be seen easily in both Figure 4 and Figure 5, which shows the PSD of the signals before and after Gabor denoising. It can be observed that only high power heart sound components were affected

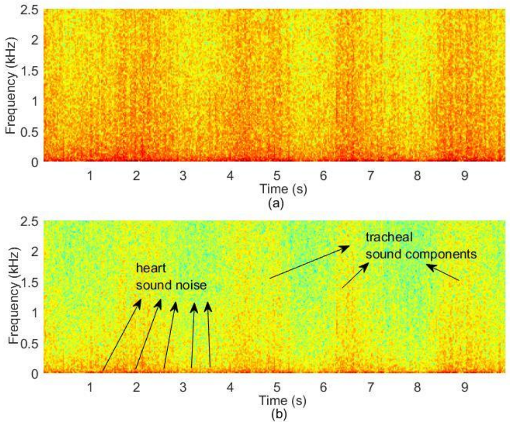


Figure 2. Typical representations of (a) raw recorded tracheal sound signal spectrogram, and (b) the spectrogram of the same signal after RLS-ANC adaptive filtering.

from the denoising procedure, leaving other parts untouched. Figure 5 also includes the magnified low frequency part of the PSD, which emphasised that the tracheal sound signal was considerably decreased at the frequencies up to 150 Hz. This clearly demonstrates the effectiveness of the masking technique.

5. Conclusions

We removed the background noise and heart sound noise in the tracheal sound signal successfully by RLS-ANC adaptive filtering, generalized biorthogonal Gabor expansion and Gabor coefficient masking method. Both heart sound signal localization and filtering were done by the Gabor expansion. The noise filtering in the biomedical signals by Gabor expansion was done in the previous studies [19]. Here we applied the Gabor expansion to tracheal sound filtering problem and achieved the noise-free traceal sound signal at the end. It is proved that the respiratory sound signal is very well modelled by Gabor coefficients. Although the respiratory sound as a time-varying signal covers very large area in the time-frequency domain, the heart sound as a noise signal has very distinctive location and can be easily processed by the Gabor expansion. The linearity of the Gabor expansion suggests the possibility of further processing of the respiratory sound signals. For instance, one may consider the cross spectral analysis between the tracheal sound signal and the lung sound signal. Furthermore, similar analysis can be carried out by the selection of the windows for the adventitious sound spectrum. Finally, comparing to our previous study [20], although the figures shows the similar results, in terms of the computational cost and simulation duration Gabor expansion technique is more superior than the spectrogram and adaptive filtering technique.

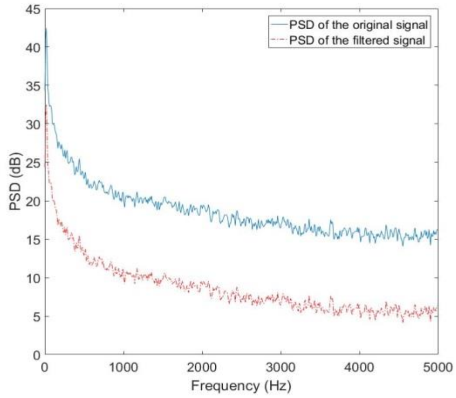


Figure 3. PSD comparisons of the tracheal sound signals before and after RLS-ANC filtering. (Solid line represents PSD of the original signal; broken line represents PSD of the broad band noise filtered signal).

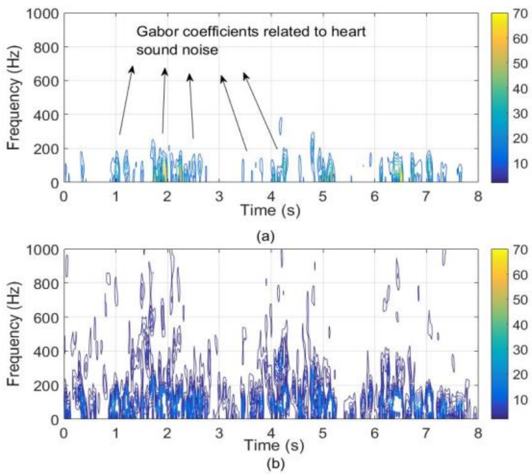


Figure 4. Typical representations of Gabor coefficients (magnitudes) for (a) the tracheal sound segment after RLS-ANC filtering and (b) same segment after soft clipping (denoising).

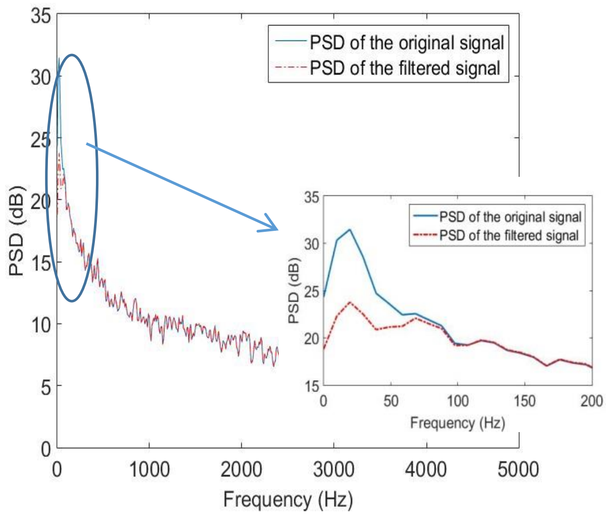


Figure 5. PSD comparisons of the tracheal sound signals before and after Gabor denoising. (Solid line represents PSD of the tracheal sound signal after RLS-ANC filtering, broken line represents PSD of the tracheal sound signal after Gabor denoising). Inserted subfigure is the magnified region between $0 - 200\ Hz$.

6. References

[1] H. Pasterkamp, S. S. Kraman, G. R. Wodicka, "Respiratory sounds advances beyond the stethoscope", *Am J Respir Crit Care Med*, vol. *156,* pp. *974-987,* 1997.

[2] Z. Moussavi, M. T. Leopando, H. Pasterkamp, G. Rempel, "Computerized acoustical respiratory phase detection without airflow measurement", *Med Biol Eng Comput*, vol. *38,* no. *2,* pp. *198-203,* 2000.

[3] A. Yadollahi, Z. Moussavi, "A robust method for estimating respiratory flow using tracheal sounds entropy", *Med Biol Eng Comput*, vol. *53,* no. *4,* pp. *662-668,* 2006.

[4] H. Zheng, H. Wang, L. Y. Wang, "Lung sound pattern analysis for anesthesia monitoring", in *Proceedings of the American Control Conference*, Portland, USA, 2005, pp. *1563-1568.*

[5] J. Gnitecki, Z. Moussavi, H. Pasterkamp, "Recursive least squares adaptive noise cancellation filtering for heart sound reduction in lung sounds recordings", in *Proceedings of the 25th Conf. Engineering in Medicine and Biology Society, EMBC03*, Cancun, Mexico, 2003, pp. *2416-2419.*

[6] V. K. Iyer, P. A. Ramamoorthy, H. Fan, Y. Ploysongsang, "Reduction of heart sounds from lung sounds by adaptive filtering", *IEEE Trans Biomed Eng*, vol. *33,* no. *12,* pp. *1141-1148,* 1986.

[7] J. Gnitecki, I. Hossain, Z. Moussavi, H. Pasterkamp, "Qualitative and quantitative evaluation of heart sound reduction from lung sound recordings", *IEEE Trans Biomed Eng*, vol. *52,* no. *10,* pp. *1788-1792,* 2005.

[8] M. T. Pourazad, Z. Moussavi, G. Thomas, "Heart sound cancelation from lung sound recordings using adaptive threshold and 2D interpolation in time-frequency domain", in *Proceedings of the 25th Conf. Engineering in Medicine and Biology Society, EMBC03*, Cancun, Mexico, 2003, pp. *2586-2589.*

[9] E. Saatci, A. Akan, "Lung sound noise reduction using Gabor time-frequency masking", in *Proceedings of the World Congress on Medical Physics and Biomedical Engineering WC*, Seoul Korea, 2006, pp. *865-868.*

[10] S. Qian, D. Chen, "Discrete Gabor transform", *IEEE Trans Signal Process*, vol. *41,* pp. *2429-2438,* 1993.

[11] J. Yao, P. Krolak, C. Steele, "The generalized Gabor transform", *IEEE Trans Image Process*, vol. *7,* pp. *978-988,* 1995.

[12] M. Bastiaans, M. C. W. Geilen, "On the discrete Gabor transform and the discrete Zak transform", *Signal Processing*, vol. *49,* pp. *151-166,* 1996.

[13] H. Rotstein, S. Raz, "Gabor transform of time-varying systems: exact representation and approximation", *IEEE Trans Auto Cont*, vol. *44,* pp. *729-741,* 1999.

[14] X. G. Xia, S. Qian, "On the rank of the discrete Gabor transform matrix", *Signal Processing*, vol. *81,* pp. *1083-1087,* 2001.

[15] R. Vio R, W. Wamsteker, "Joint time-frequency analysis: a tool for exploratory analysis and filtering of non-stationary time series", *Astronomy & Astrophysics*, vol. *388,* no. *3,* pp. *1124-1138,* 2002.

[16] D. L. Donoho, "Denoising via soft thresholding", *IEEE Trans Inform Theory*, vol. *41,* no. *3,* pp. *613-627,* 1995.

[17] P. J. Wolfe, S. J.Godsill, W. J. Ng, "Bayasian variable selection and regularization for time-frequency surface estimation", *J of the Royal Stat Soc*, vol. *66,* no. *3,* pp. *575-589,* 2004.

[18] S. Haykin, "Adaptive Filter Theory", Prentice Hall, 1991.

[19] Y. Lu, S. Joshi, J. M. Morris, "Noise reduction for NMR FID signals via Gabor expansion", *IEEE Trans Biomed Eng*, vol. *44,* pp. *512-528,* 1997.

[20] E. Saatci, A. Akan, "Heart sound reduction in lung sounds by Spectrogram", in *Proceedings of the 3rd European Medical and Biological Engineering Conference, Prague*, Seoul Korea, 2005, pp. *11(1).*

Esra Saatçı received the B.Sc. (1993) and Ph.D. (2009) degrees in electronic and biomedical engineering from Istanbul University, Turkey respectively; and M.Sc. (1995) degree in biomedical engineering from University of Surrey, UK. She worked for five years at respiratory and anesthesia instrumentation companies between 1995 and 1999 in Istanbul. She was a research assistant in mechanical engineering department at Kings College London between 2000 and 2004; and in electrical & electronics engineering department at Istanbul Kültür University between 2004 and 2009. Since 2009, she is an Assitant Professor at the Deparment of Electrical & Electronics Engineering, Istanbul Kültür University. Her current research interests include statistical signal processing with applications of biomedical signals, respiratory models and respiratory signal processing.

Aydin Akan received the B.Sc. degree from the University of Uludag, Bursa, in 1988, the M.Sc. degree from the Technical University of Istanbul, Istanbul, Turkey in 1991, and the Ph.D. degree from the University of Pittsburgh, Pittsburgh, PA, USA, in 1996, all in electrical engineering. He was with the Department of Electrical and Electronics Engineering, University of Istanbul, between1996 and 2017. He is currently a Professor at the Department of Biomedical Engineering, Izmir Katip Celebi University, Cigli, Izmir. His current research interests include nonstationary signal processing, time–frequency signal analysis methods and their applications to wireless communications and biomedical engineering. He is a senior member of the IEEE Signal Processing Society and an Associate Editor of the Digital Signal Processing Journal.

IMPROVEMENT OF MAGNETIC FILTERS' PERFORMANCE BY CONTROLLING REGIONAL FIELD WITH PWM USING A DIGITAL SIGNAL CONTROLLER

Ömerül Faruk ÖZGÜVEN[1]

[1]Inonu University, Engineering Faculty, Biomedical Engineering, Malatya
omer.ozguven@inonu.edu.tr

Abstract: *In this paper, a new method for the improvement of the performance of magnetic filters for the removal of disperses mixtures with magnetic characteristics from industrial liquids and gases is investigated. In order to accelerate the reduction of the concentration of dispersed mixture, it is suggested that the intensity of the external magnetic field throughout the magnetic filter should be adjusted regionally. To achieve this goal, an intermediate control circuit with PWM (Pulse Width Modulation) driver capable of driving the external magnetic fields of magnetic filter regionally based on three regions is designed. The dsPIC30F2010 Digital Signal Controller (DSC; Microchip®) is used in the controller circuit. The experimental results show that the concentration of dispersed magnetic mixture contained in the aqueous suspension passed from magnetic filter is reduced more efficiently. It is claimed that due to the adjustment of the external magnetic field intensity applied to the magnetic filters throughout the magnetic filter, the filter performance is increased, the electrical energy consumption is reduced, and the optimum design of magnetic filters is achieved.*
Keywords: *Magnetic filters, disperse particle concentration, control circuits, Digital Signal Controller (DSC), dsPIC30F2010, Pulse Width Modulation (PWM)*

1. Introduction

The removal process of magnetic-featured dispersed mixtures from liquids and gases in the external magnetic fields have been known since ancient times [1]. The theory and the implementation of cleaning these environments from low-concentration and micron-sized dispersed mixtures in high gradient magnetic fields has rapidly evolved in the past four decades [1-4]. The methods called High Gradient Magnetic Separation (HGMS) and High Gradient Magnetic Filtration (HGMF) have been used successfully for the enrichment process of minerals in the mining engineering [1] and applied to various problems in environmental engineering as an effective method [5-7].

Nowadays, HGMF methods have been used in various technological processes such as energy production, nuclear, chemical and petroleum industry. They provide increased efficiency in these industrial areas [8-11].

In recent years, systems and devices primarily developed for high gradient magnetic fields have been adopted in the magnetophoresis operations of nano and micro technological processes widely [10,12].

The principals of HGMS and HGMF have been used effectively in the processes of magnetic drug targeting, immunomagnetic cell separation, purification of fermentation products and purification of recombinants in medical and biological systems [13-17].

High gradient magnetic systems have also been suggested as advantageous methods in many other technological processes [18-22].

Various high-gradient magnetic field structures have been used in the systems of HGMS, HGMF and other applications [23, 24]. The reason for using all these systems in industrial and other applications is due to their very high performance. Therefore high-gradient magnetic field systems and devices require an optimum construction and high performance.

In many studies, the superconductive magnetic systems have been used as external magnetic field sources in order to increase the effectiveness of high gradient magnetic fields systems [11, 25, 26].

Received on: 07.02.2017
Accepted on: 13.03.2017

Numerous studies have been carried out to determine optimal system parameters by improving the theory of these systems and by increasing performances of HGMS and HGMF systems [2, 4, 7, 27-32].

In all these theoretical and practical studies, the parameters of both geometric and operation parameters such as concentration of dispersed mixtures, flow rate, external magnetic field intensity, temperature, and filter length etc. of high gradient magnetic field systems have been assumed to be constant. However the reference values of these parameters vary due to operation conditions. These conditions substantially affect the system performance.

During the operation of these systems, whether the system performance should be kept constant or should be improved is one of the current problems from the practical point of view. For this purpose, during the operation of high-gradient magnetic field systems, it is necessary to control the basic parameters and to make the required adjustments. It is not possible to carry out this operation neither by means of the theoretical and experimental models mentioned above nor by others proposed in the related literature.

In this study, increased capture performance of particles by magnetic filter from liquid and gases including low concentration and fine sizes particles with magnetic properties are proposed. For that purpose as a new method, regional regulation method used to determine the values of magnetic field of filter's magnetic system.

By investigating the parameters affecting the performance of HGMF, the current value in the winding of external magnetic field system (solenoid) is adjusted regionally throughout the filter. In accordance with this purpose, a PWM driver control system [33,34], which can adjust the current value of model filter solenoid system within three regions with time, is designed and the effects of this system on the filter performance is investigated. Stable operation conditions of the designed control system have been examined and their characteristics have been determined. Some suggestions have been proposed to operate this control system with HMGF used in industry.

2. Formulation of the problem

As shown in Figure 1, the magnetic filter or separator consists of non-magnetic body, magnetic system generating an external magnetic field (magnet, solenoid, core-type electromagnet), and input and output pipes and filter matrixes.

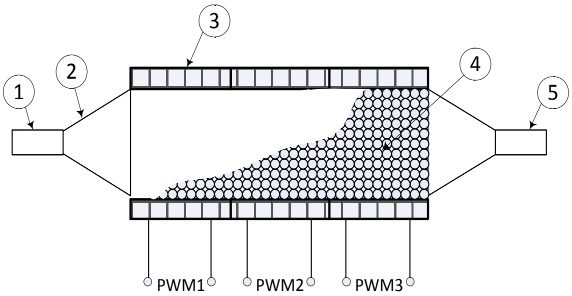


Figure 1. The principal scheme of the magnetic filter, 1- The inlet pipe, 2- The body, 3- The external electromagnetic field system (solenoid) 4- The filter matrix, 5- The outlet pipe.

The matrix elements of magnetic filter consist of magnetisable ferromagnetic materials (spheres, rods, wires, metallic turnings). These matrix elements can easily be magnetized with the effect of external magnetic field and thus they constitute local high gradient magnetic fields around themselves and at tangent points. While the suspensions including micron and submicron-sized particles are passing from matrix, the particles held by high gradient magnetic fields are collected in this region [9]. The efficiency or performance of magnetic filter is determined by the cleaning coefficient [8,9]:

$$\frac{\psi}{\lambda} = 1 - \frac{C_{out}}{C_{in}} \tag{1}$$

Where ψ is the cleaning coefficient of the filter, λ is the rate of the magnetisable particles of the mixture within suspension, C_{in} and C_{out} are the filter inlet and outlet concentrations of the mixture within suspension, respectively.

In general, the filter performance depends on geometric, magnetic, hydrodynamic and rheological parameters of the system. The most effective parameters from these parameters are the length of the magnetic filter (L), the filtration velocity (V_f) and the external magnetic field intensity (H or B) [27-29]. The magnetic filter performance in non-stationary state can be expressed as follows [9]:

$$\psi = \lambda[1 - e^{(\alpha t/t_0 - \beta x)}] \tag{2}$$

Where t is the time, t_0 is the stable operation time of the filter, x is the length through the filter, β is the adsorption coefficient of particles, α is the detachment coefficient of the adsorbed particles.

It is determined that the operation parameters of the real magnetic filters used both in industry and in laboratory is around at $V_f \leq 0.1$ m/s, $H \leq 150$ kA/m, β=(1…2) m^{-1}, α=0.1-0.6 h^{-1}, L=1-1.2 m [7, 8]. The particle adsorption coefficient of filter matrix (β), constituted from ferromagnetic spheres can be presented as follows [9, 30, 31]:

$$\beta = 2.6\ 10^{-3} \left(\frac{V_m}{V_f}\right) \frac{1}{d} \tag{3}$$

$$V_m = \frac{3k\mu_0 \mu^{1.38} H^2 \delta^2}{\eta d} \tag{4}$$

Where, k is the magnetic susceptibility of adsorbed particle, μ is the magnetic permeability of matrix elements, δ is the particle size (diameter), d- is the sphere diameter of matrix element, η is the dynamic viscosity of the suspension.

As can be seen, one of the most important parameters affecting the performance of the magnetic filter is the external magnetic field intensity (H). While the magnetic filter is in operation, an adjustment on the magnetic field intensity (H) causes a change in the filter performance. In practice, the change of magnetic field intensity varies in a wide range H = (10-150) kA/m [8, 9]. At low and average values of the magnetic field intensity (H = (10-100) kA/m), to obtain the filter cleaning coefficients ψ = % (70…85), the filter length must be around L = (1-1.2) m [8]. Experimental studies have shown that it is possible to obtain high performance from the magnetic filter even at low magnetic field values [8,9]. In this case, particles accumulated in the filter are aggregated in a certain area on the input side of the filter. Therefore, the remaining region of the active length of the filter operates at very low efficiency. Not using the filter length optimally results in both the rise in the amount of consumed electricity and the waste of the magnetic materials over time; thus, leading to the reduction in the filter performance. Accordingly, when the magnetic filter is in operation, various regions of the filter matrix could be run in different operation states by means of creating non-homogeneous distribution of the magnetic field intensity throughout the filter. It is possible to adjust all regions of the filter matrix automatically with the same performance by taking care of operation conditions of the filter system. In this case, the supply currents of winding or coils providing the external magnetic field of the magnetic filter must be adjusted locally. While coils in the input region of the magnetic filter are supplied with less current, the currents of windings towards the filter output must be automatically and regionally increased. In this study, an automatically controlled PWM drive circuit is proposed to achieve the goal of driving the magnetic filter windings accordingly

3. Material and method

3.1. The characteristics of the magnetic filter

Solenoid type magnetic filter windings with the total length of L_1 = 300 mm, L_2=300 mm and L_2=250 mm and with the inner diameter of 25 mm, consist of three regions with $l_1 = l_2 = l_3$ = 100 mm. The ferromagnetic spheres whose diameters are d = 5 mm are used as the filter matrix element. Magnetic filter and filter matrix element used in experimental are shown in Figure 2 and Figure 3, respectively.

The aqueous solution, whose size is δ = 1 - 2 μm, is used as the cleaning environment. The previously prepared magnetic section containing corrosion particles is around at % 80-85, (λ = 0.8-0.85).

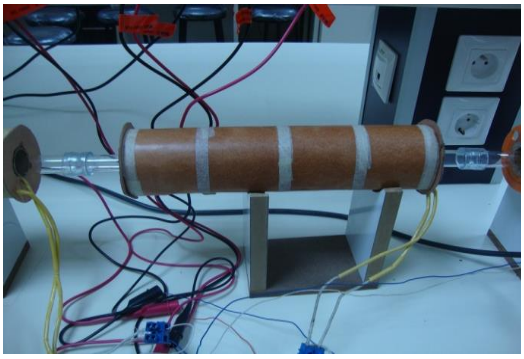


Figure 2. Solenoid type magnetic filter

Figure 3. Filter matrix

Flow of aqueous suspension from filter matrix has been carried out as non-recycled (single-layer) and the filter speed has been around V_f = 0.01m/s.

The voltage applied to the windings of the solenoid could be adjusted in the range of U = 0V-50V by changing the duty ratio of the PWM signal. In this case, the current in solenoid windings was limited at $I \leq 2.5$ A and the magnetic field intensity in the solenoid centre was measured as $H \leq 30$ kA/m. Both the control circuit and the software have been realized for driving solenoid windings by means of PWM signals, which provide regional (three regions) max 50V DC voltage. Due to its low cost and good technical properties, the dsPIC30F2010 Digital Signal Controller was used in the control circuit.

3.2. PWM control scheme

PWM control scheme, producing the required PWM signals and adjusting duty ratios by using the dsPIC30F2010 Digital Signal Controller (DSC), is shown in Figure 4. The dsPIC30F2010 DSC contains 16x16-bit working register array, 10-bit Analog-to-Digital Converter (A/D) with six input channels and three PWM modules [35]. Also, the DSC can multiply and divide signed or unsigned fractional/integer numbers in machine language. Because of these advantages, the control program was written in machine language. In terms of performance, the dsPIC30F2010 has superior properties than 16F and 18F series processors previously manufactured by Microchip [36].

In Figure 4, three potentiometers are connected to AN0, AN1 and AN2 analog inputs of the dsPIC30F2010. These potentiometers adjust the duty rates of FILTER1, FILTER2 and FILTER3, respectively. The adjusted duty rates are

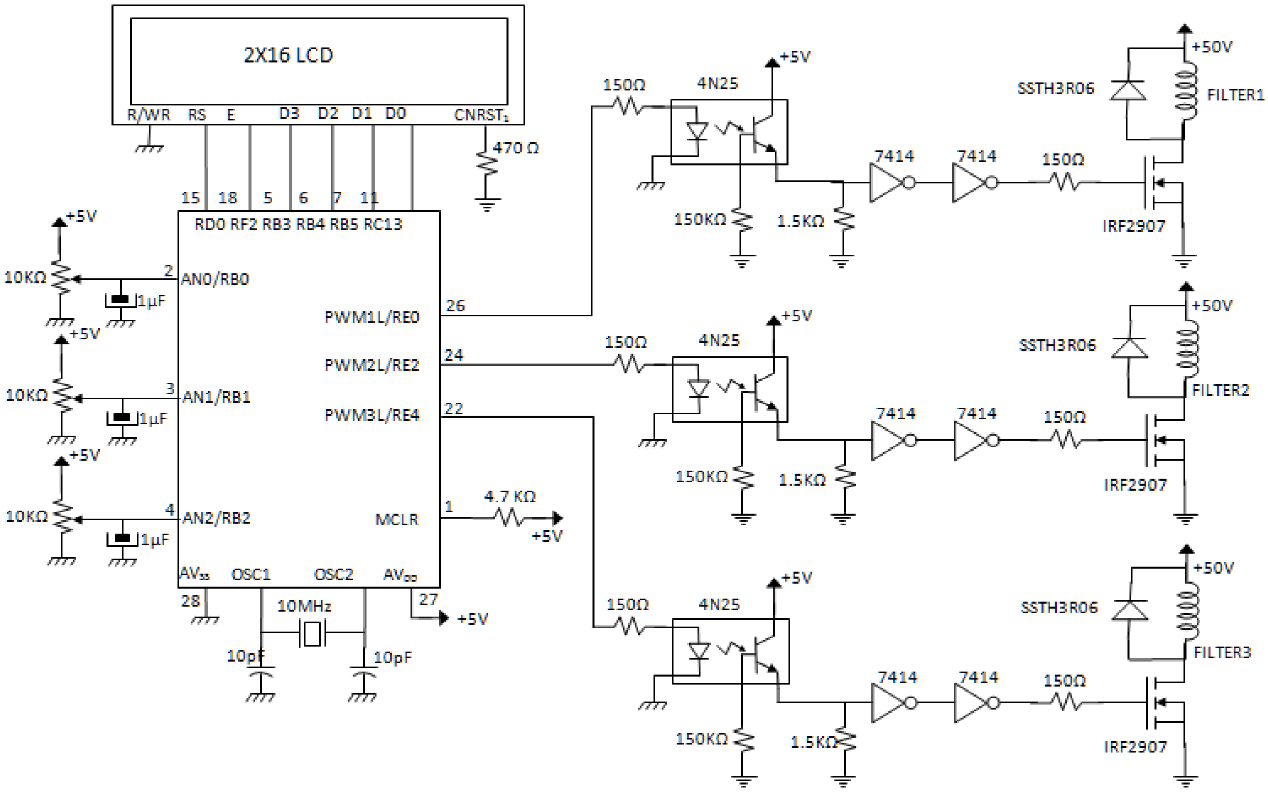


Figure 4. PWM Control Scheme using dsPIC30F2010

shown on an LCD as percentages under labels of DUTY1, DUTY2 and DUTY3. The PWM output signals are taken from PWM1L, PWM2L and PWM3L pins of the DSC and are applied to the driver inputs. IRF2907 MOSFET components and opto-couplers are used to convey the PWM signals to solenoids in the drive circuit. The drain-source voltage and drain current of IRF2907 are 75V (V_{DSS}=75V) and 75A (I_D=75A), respectively.

3.3. The flowchart of the control program

Figure 5 shows the flowchart of the control program where the initialization of peripheral devices of dsPIC30F2010 DSC is achieved and three separate PWM signals are generated.

In Figure 5. (a), Stack Pointer, ADC and PWM module, PORTE and LCD are initialized. After writing DUTY1, DUTY2 and DUTY3 messages on LCD, the interrupt request initialization is done for ADC. Then, the Interrupt request is enabled and the main program waits for the ADC interrupt request.

The flowchart of the ADC interrupt request program is shown in Figure 5. (b). Firstly, the digital value corresponding to the analog input for each channel is read. These digital values determine the duty ratio of PWM signal to be applied to solenoids. The values read from ADC are firstly scaled with the equation (5). Then they are loaded to PDC [35] determining the duty ratio in the processor:

$$PDC = \frac{PDC_{max}}{1023} X_{ADC} \quad (5)$$

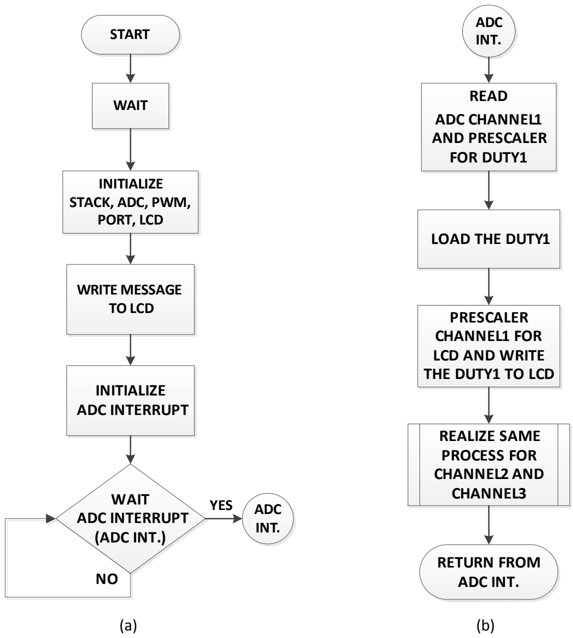


Figure 5. The flowchart of a) The initialization program, b) The interrupt request program.

Where X_{ADC} refers to the digital value of the analog input read from the related channel of ADC. This digital value has the maximum value of 1023.

As f_{PWM} is 1 KHz and $T_{DUTYmax}$ is 1ms, the maximum value to be loaded to PDC is 5000 (PDC_{max} = 5000).

The scaling shown in equation (6) is applied to show the duty ratio of each PWM signal on LCD as a percentage from % 0 to % 100.

$$\%DUTY = \frac{100}{1023} X_{ADC} \qquad (6)$$

The scaling operations shown in equations (5) and (6) can easily be implemented in assembly language, thanks to the 16 bit multiplication and division instructions of the dsPIC30F2010 processor. The control program is shown in appendix.

The experimental setup and output signal on filter2 are shown in Figure 6, Figure 7, respectively.

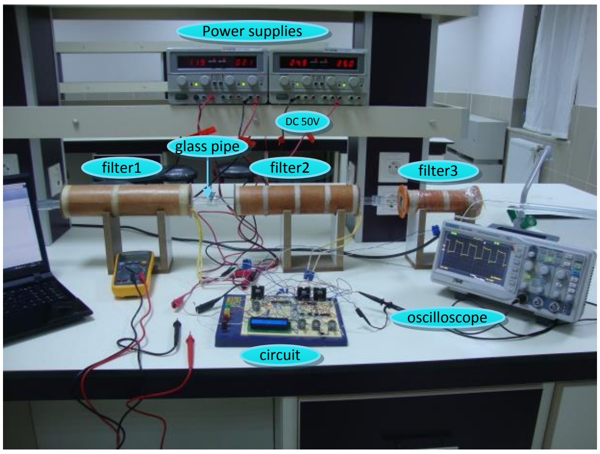


Figure 6. The experimental setup

Figure 7. The PWM output signal on filter2 (*f*=1 KHz, Duty=%50, volt/div x10)

4. Results and discussions

The results of [9] are taken into account to assess the effect of splitting the electromagnetic solenoid into two or more regions and then by applying different voltage values to each region. Then equations (2), (4) are used in order to define the experimental concentration output variation of these results. Each solenoid used in this study consists of three independent solenoids of which length is 100 mm and the total length of the solenoid is 300 mm. Each solenoid is driven by a PWM driver output in the control circuit, separately. By adjusting duty rates of PWM signals by means of three separate potentiometers, individual solenoid currents were adjusted in the range of 0.1A -0.8A. Ferromagnetic spheres in 5 mm diameter were used as magnetic filter matrix elements. The suspension prepared as C_{in}=50 ppm has been passed from the magnetic filter as non-recycled. The filter performance or concentration rates calculated by determining the magnetic filter output concentration of the magnetic dispersion mixture in suspension.

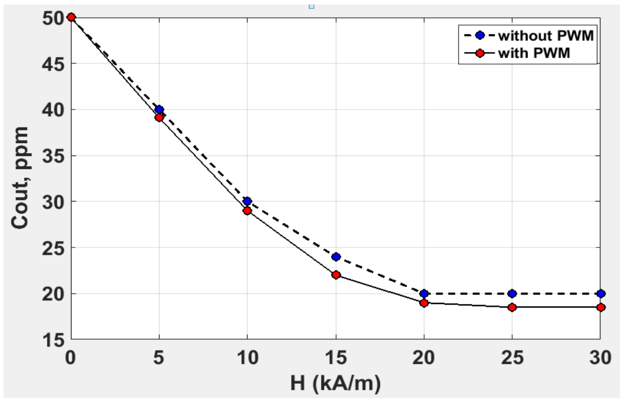


Figure 8. The change of output concentration of captured particles relationship with external magnetic field intensity.

The changes of concentration and effectiveness coefficients of magnetic filter respectively are demonstrated with PWM and without PWM situations at Figure 8 and Figure 9.

The various of external magnetic field intensity are adjusted for first solenoid accordingly Figure 8 and Figure 9. The magnetic field of second and third intensity solenoid is respectively adjusted based on equations (1) and (2) by using PWM control system. When the PWM is used, the filter performance will be increased as shown in both Figure 8-9. In this case, results will be more effective in magnetic filters and separator which have a high flow rate used PWM.

Accordingly filter length considering equation 1-3, the change of magnetic filter performance and concentration ratio can be evaluated. This comparison is shown at Figure 10. By adjusting duty rates of PWM signals by means of three separate potentiometers, individual solenoid currents were respectively selected as 0.510A, 0.625A and 0.765A in the flow direction of the suspension in the magnetic filter.

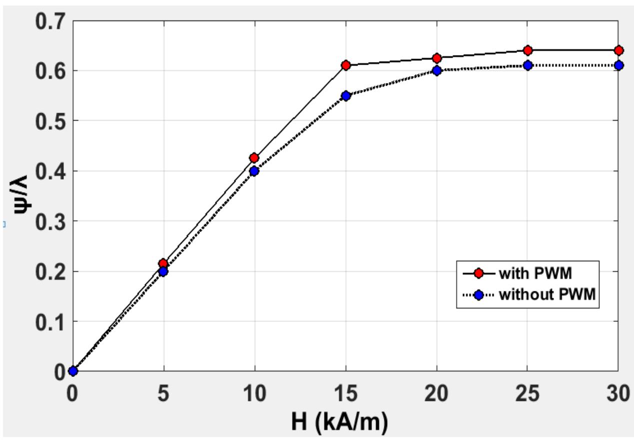


Figure 9. The change of effectiveness coefficient (performance) of the filter relationship with external magnetic field intensity.

For these current values, the magnetic flux density (the magnetic field intensity) in the center of the solenoid with no core are as follows B(H)=0.019T (1520 A/m), 0.022 T (17507 A/m) and 0.028 T (22282 A/m),respectively.

As can be seen from Figure 10, the concentration in the filter output of the magnetic disperses mixture due to the use of the PWM control circuit, is less than the one without PWM. This is due to the fact that relatively small size and weak magnetic-featured particles are hold in the regions with higher magnetic field instead of the entrance of the filter. Therefore, because of the regional non-homogeneous magnetization of the filter matrix, the possibility of the filter performance increases. This result is of great importance in the magnetic filter and separators with ferromagnetic matrix. Because in many industries such as nuclear power plants, chemical technology, paper industry, the sizes and concentrations of disperse mixtures within fluids used, are very low. In the technological processes using these fluids, high quality cleansing is necessary for the fluid environments from disperse mixtures. In this case, as proposed in this paper the use of PWM control units in electrical circuits of the magnetic filters and separators may be the most effective method. Furthermore, when driving the magnetic filters with the conventional techniques, solenoids are driven directly without PWM. This means that the solenoids are supplied with high electrical current consistently. In addition, driving solenoids of magnetic filters with previously known conventional methods increases the complexity and the cost of the electronic circuit design and thus the realization of such circuits becomes very difficult. As a result, power elements used in classical filters draw more current and therefore consume more electric power.

Driving the magnetic filter with the PWM technique, namely switching the current of the solenoid on and off to obtain the desired amount in a fixed period, reduces the power consumption and enables the solenoid to be driven easily.

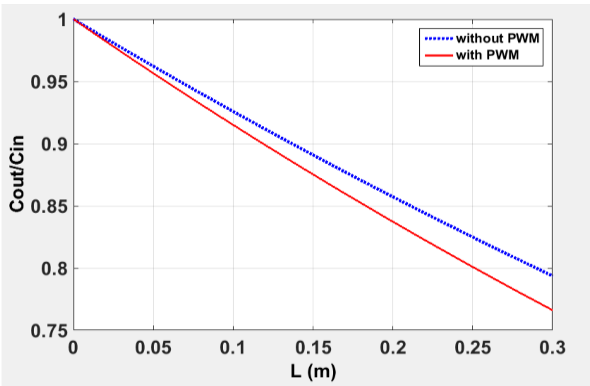


Figure 10. The change of filter performance for filter lengths.

This method also allows supplying solenoid with the optimum regional current value in a desired period. Because of long operation time of magnetic filters and separators, it is clearly seen that driving solenoids with PWM technique provides a significant contribution to electrical energy saving. Thus, using of PWM techniques and control circuits in the magnetic filters and separators will possibility provide an improvement in the performance of these systems, facilitate the design of control circuit, reduce the circuit cost and power consumption, and increase the effectiveness of the technological processes.

5. Conclusions

In order to increase the performance of magnetic filters and separators it is possible to regionally adjust magnetic fields of these devices. For this purpose, multi-channel PWM drive circuits can be used. Magnetic filters can be driven with multi-regional adjustable voltage by increasing the number of magnetic filter windings and with the help of specially designed PWM drivers. In this case, the performance of magnetic filter can be increased by using solenoids with different magnetic field intensities in different regions.

The PWM control circuit design proposed in this paper allows driving magnetic filters with 300 W (by using IRF2907) or higher power. This situation allows the use of PWM control in industrial type electrical circuits of magnetic filters and separators effectively. The PWM control circuits not only permit the adjustment of the current and voltage of magnetic circuits of magnetic filters and separators, but also it provides the optimal energy savings for these systems. The use of PWM driver circuits in the magnetic filters reduces cost of circuit and facilitates the circuit design. Connecting PWM control circuits in serial and parallel can be effectively used in the operation of magnetic filters and separators.

6. References

[1] J. Svoboda, “Magnetic Techniques for the Treatment of Minerals”, Kluwer Academic Publishers: Dordrecht, the Netherlands, 2004.

[2] J.H.P. Watson, “Magnetic Filtration”, *J. Appl. Phys.*, vol. *44*, no. *9*, pp. *4209-4213*, 1973.

[3] F. J. Friedlaender, R. R. Gerber, H.P., Henkel, et al., “Particle Build up on Single Spheres in HGMS”, *IEEE Trans. Magn.*, vol. *17*, no. *6*, pp. *2804-2806*, 1981.

[4] R. Gerber, R.R. Birss, “High Gradient Magnetic Separation”, John Wiley and Sons, New York, 1983.

[5] C. *De Latour*, “Magnetic separation in water pollution control”, *IEEE Trans. Magn.* vol. *9*, no. *3*, pp. *314-316*, 1973.

[6] C. Tsouris, J. Noonan, T.Y. Ying, S. Yiacomi, “Surfactant effects on the mechanism of particle capture in high-gradient magnetic filtration”, *Sep. Purify. Technol.* vol. *51*, no. *2*, pp. *201–209*, 2006.

[7] Sh.-J. Bai, Sh.-M. Wen, D.-W. Liu, W.-B. Zhang, “Separation of phosphorous and magnetic mineral fines from siderite reductive ore by applying magnetic flocculation”, *Separation Science and Technology*, vol. *49*, no. *9*, pp. *1434-1441*, 2014.

[8] A.V. Sandulyak, “Magnetic Filtration of Liquids and Gases”, Chemics, Ed. Moscow, 1998.

[9] T. Abbasov, “Electromagnetic filtration methods (Theory, application and construction)”, Seçkin Ed. Ankara, 2002.

[10] L.E. Udreal, N.J.C. Strachan, V.B. Badescu, O. Rotariu, “An in vitro study of magnetic particle targeting in small blood

vessels", *Physics in Medicine & Biology*, vol. *51*, pp. *4869–4881*, 2006.
[11] N. Mizuno, F. Mishima, Y. Akiyama, H. Okada, N. Hirota, H. Matsuura, T. Maeda, N. Shigemoto, S. Nishijima, "Removal of Iron Oxide With Superconducting Magnet High Gradient Magnetic Separation From Feed-Water in Thermal Plant", *IEEE Transactions on Applied Superconductivity*, vol. 25, no. 3, 2015.
[12] N. Pamme, "On-chip bioanalysis with magnetic particles", *Current Opinion in Chemical Biology*, vol. *16*, no . *3-4*, pp. *436-443*, 2012
[13] E.P. Furlani, X. Xue, "Field force and transport analysis for magnetic particle-based gene delivery", *Microfluids and Nanofluids*, vol. *13*, no. *4*, pp. *589-602*, 2012.
[14] M. Zborowski, J.J. Chalmers, "Magnetic Cell Separation", Elsevier, Amsterdam, the Netherland, 2008.
[15] I. Safarik, M. Safarikova, "Magnetic techniques for the isolation and purification of proteins and peptides", *Biomagn. Res.Technol*., vol. *2*, no. *7*, 2004.
[16] P. Fraga Garcia, M. Brammen, M. Wolf, S. Reinlein, M. Freiherr von Roman, S. Berensmeier, "High-gradient magnetic separation for technical scale protein recovery using low cost magnetic nanoparticles", *Separation and Purification Technology*, vol. *150*, pp: *29–36*, 2015
[17] A.Pasteur, N. Tippkötter, P. Kampeis, and R. Ulber, "Optimization of High Gradient Magnetic Separation Filter Units for the Purification of Fermentation Products", *IEEE Transactions on Magnetics*, vol. 50, no. 10, 2014.
[18] F. Larachi, D. Desvigine, "Ferrofluid induced-field effects in homogeneous porous media under linear-gradient dc magnetic fields", *Chemical Engineering and Processing*, vol. *46*, no. *8*, pp. *729-735*, 2007.
[19] J. Hristov, "Magnetic Field assisted fluidization – a unified approach. Part 7. Mass transfer: Chemical reactors, basic studies and practical implementations thereof", *Reviews in Chemical Engineering*, vol. *25*, no. *1-3*, pp. *1-254*, 2009.
[20] J. Hristov, "Magnetic assisted fluidization – a unified approach. Part 8. Mass transfer: Magnetically assisted bioprocesses", *Reviews in Chemical Engineering*, vol. *26*, no. *3-4*, pp. *55-128*, 2010.
[21] J. Hristov, "Magnetic field assisted fluidization - a unified approach. Part 9. Mechanical processing with emphasis on separations", *Reviews in Chemical Engineering*, vol. *28*, no. *4-6*, pp. *243-308*, 2012.
[22] J. Hristov, L. Fachikov, "An overview of separation by magnetically stabilized beds: State of the art and potential application", *China Particuology*, vol. *5*, no. *1-2*, pp. *11-18*, 2007.
[23] T. Karadag, Z. Yildiz, A. Sarimeseli, T. Abbasov, "Estimation of Magnetization Properties of the Ferromagnetic Poly-Granular Beds", *Journal of Dispersion Science and Technology*, vol. *31*, no. *6*, pp. *826-830*, 2010
[24] B. A. Guseva, V. G. Semenovb, V. V. Panchuk, "Numerical Simulation of High-Gradient Magnetic Filtration", Technical Physics, Vol. 61, No. 9, pp. 1292–1298, 2016.
[25] A. Eskandarpour, K. Iwai, Sh. Asai, "Superconducting Magnetic Filter: Performance, Recovery, and Design*",* *IEEE Transactions on Applied Superconductivity*, vol. 19, no. 2, pp. 84-95, 2009.
[26] H. Okada, K. Imamura, N. Hirota, T. Ando, S. Shibatani, N. Mizuno, M. Nakanishi, F. Mishima, Y. Akiyama, S. Nishijima, H. Matsuura, T. Maeda, N. Shigemoto," Development of a Magnetic Separation System of Boiler Feedwater Scale in Thermal Power Plants", *IEEE Transactions on Applied Superconductivity*, vol. 26, no. 3, 2016.
[27] J.H.P Watson, "Approximate solutions of the magnetic separator equations", *IEEE Trans. Magn*., vol. *14*, no. *4*, pp. *240-245*, 1978.
[28] S. Uchiyama, S. Kondo, M. Takayasu, I. Eguchi, "Performance of parallel stream type magnetic filter for HGMS", *IEEE Trans. Magn*. Vol. *12*, no. *6*, pp. *895-897*, 1976.
[29] N. Rezlescu, V. Murariu, O. Rotariu, V. Badescu, "Capture modeling for an axial high gradient magnetic separation filter with a bounded flow field", *Powder Technology*, vol. *83*, no. 3, pp. *259-264*, 1995.
[30] T. Abbasov, M. Köksal, S. Herdem, "Theory of high gradient magnetic filter performance", *IEEE Trans. Magn.*, vol. *35*, no. *4*, pp. *2128-2132*,1999.
[31] T. Abbasov, S. Herdem, M. Köksal, , "Performance of high gradient Magnetic filters with granular Matrix", *Separation Science and Technology*, vol. *34*, no. 2, pp. *263-276*, 1999.
[32] S. N. Podoynitsyn, O. N. Sorokinan, A. L. Kovarski, "High-gradient magnetic separation using ferromagnetic membrane", *Journal of Magnetism and Magnetic Materials*, vol. *397*, pp: *51–56*, 2016
[33] M. Tümay, V. Karslı, H.F. Aksoy, "Computer simulation of three phase electrical machines and adjustable-speed AC drives", *Computers and Electrical Engineering*, vol. *28*, pp. *611–629*, 2002.
[34] C. A. Gonzalez-Gutierrez, J. Rodriguez-Resendiz, G. Mota-Valtierra, et. al., "PC-based architecture for parameter analysis of vector-controlled induction motor drive", *Computers and Electrical Engineering*, vol. *37*, pp. *858–868*, 2011.
[35] dsPIC30F Family Reference Manual, Available: http://ww1.microchip.com/downloads/en/DeviceDoc/70046e.pdf (Accessed in 2017)
[36] dsPIC® Digital Signal Controllers, Available: ww1.microchip.com/downloads/en/DeviceDoc/DS-70095K.pdf (Accessed in 2017)

7. Appendix

PWM Control program for dsPIC30F2010 in assembly langauge

```
;       .equ __30F2010,1
        .include "P30F2010.inc"
        .global __ADCInterrupt
        .weak   __reset
__reset:
        .equ    SAYI1,0x0800
        .equ    LW1,0x0802
        repeat  #0x3FF
        nop
        repeat  #0x3FF
        nop
        mov     #0x0900,w15
        call    PWMSRT
        call    DSPSRT
        call    LCDSRT
        call    PWMYAZ
DNG:    goto    DNG
__ADCInterrupt:
        push.s
        bclr    IEC0,#ADIE
        mov     ADCBUF0,w2
        mov     #5000,w7
        mul.uu  w2,w7,w10
        mov     #1023,w8
        repeat  #17
        div.ud  w10,w8
        mov     w0,PDC1
        mov     ADCBUF0,w2
        mov     #99,w7
        mul.uu  w2,w7,w10
        mov     #1023,w8
        repeat  #17
        div.ud  w10,w8
        mov     w0,SAYI1
        mov.b   #0x07,w9
        call    BINBCD
        mov     LW1,w11
        call    EMIR
        mov     #0xC2,w1
        call    LCDYAZ
        call    LHZR
        mov     ADCBUF1,w2
        mov     #5000,w7
        mul.uu  w2,w7,w10
        mov     #1023,w8
        repeat  #17
        div.ud  w10,w8
        mov     w0,PDC2
        mov     ADCBUF1,W2
        mov     #99,w7
        mul.uu  w2,w7,w10
        mov     #1023,w8
        repeat  #17
        div.ud  w10,w8
        mov     w0,SAYI1
        mov     #0x07,w9
        call    BINBCD
        mov     LW1,w11
        call    EMIR
        mov     #0xC7,w1
        call    LCDYAZ
        call    LHZR
        mov     ADCBUF2,w2
        mov     #5000,w7
        mul.uu  w2,w7,w10
        mov     #1023,w8
        repeat  #17
        div.ud  w10,w8
        mov     w0,PDC3
        mov     ADCBUF2,W2
        mov     #99,w7
        mul.uu  w2,w7,w10
        mov     #1023,w8
        repeat  #17
        div.ud  w10,w8
        mov     w0,SAYI1
        mov     #0x07,w9
        call    BINBCD
        mov     LW1,w11
        call    EMIR
        mov     #0xCC,w1
        call    LCDYAZ
        call    LHZR
GEC:    bset    IEC0,#ADIE
        bclr    IFS0,#ADIF
        pop.s
        retfie
PWMSRT:
        mov     #0x0000,W3
        mov     W3,PTCON
        mov     #0x0FFF,W3
        mov     W3,PWMCON1
        clr     PTMR
        mov     #2500,W3
        mov     W3,PTPER
        mov     #0x0000,W3
        mov     W3,PDC1
        mov     W3,PDC2
        mov     W3,PDC3
        bset    PTCON,#PTEN
        return
PWMYAZ:
        mov     #0x81,w1
        call    LCDYAZ
        call    GOST
        mov     #'P',w1
        call    LCDYAZ
        mov     #'W',w1
        call    LCDYAZ
        mov     #'M',w1
        call    LCDYAZ
        mov     #'1',w1
        call    LCDYAZ
        call    EMIR
        mov     #0x86,w1
        call    LCDYAZ
        call    GOST
        mov     #'P',w1
        call    LCDYAZ
        mov     #'W',w1
        call    LCDYAZ
        mov     #'M',w1
        call    LCDYAZ
        mov     #'2',w1
        call    LCDYAZ
        call    EMIR
        mov     #0x8B,w1
        call    LCDYAZ
        call    GOST
        mov     #'P',w1
        call    LCDYAZ
        mov     #'W',w1
        call    LCDYAZ
        mov     #'M',w1
        call    LCDYAZ
        mov     #'3',w1
        call    LCDYAZ
        call    EMIR
        mov     #0xC1,w1
        call    LCDYAZ
        call    GOST
        mov     #'%',w1
        call    LCDYAZ
        call    EMIR
        mov     #0xC6,w1
        call    LCDYAZ
        call    GOST
        mov     #'%',w1
        call    LCDYAZ
        call    EMIR
        mov     #0xCB,w1
        call    LCDYAZ
        call    GOST
        mov     #'%',w1
        call    LCDYAZ
        bset    IEC0,#ADIE
        bset    ADCON1,#ADON
        return
EMIR:   bclr    PORTD,#0
        repeat  #0x3FF
        nop
        nop
        return
GOST:   bset    PORTD,#0
        repeat  #0x3FF
        nop
        nop
        return
LCDYAZ:
        swap.b  w1
        mov.b   w1,w0
        call    YAZ
        call    ONAY
        swap.b  w1
        mov.b   w1,w0
        call    YAZ
        call    ONAY
        return
YAZ:    sl      W0,#0x03,W0
        mov     W0,PORTB
        bset    PORTC,#13
        btss    W0,#6
        bclr    PORTC,#13
        return
ONAY:   do      #10,KMRS
        nop
KMRS:   nop
        bset    PORTF,#2
        do      #500,KMRS1
        nop
KMRS1:  nop
        bclr    PORTF,#2
        repeat  #0x3E8
        nop
        return
DSPSRT:
        mov     #0xFFF8,w3
        mov     w3,ADPCFG
        mov     #0x00E4,w1
        mov     w1,ADCON1
        mov     #0x0408,w1
        mov     w1,ADCON2
        mov     #0x1008,w1
        mov     w1,ADCON3
        clr     ADCHS
;reading AN0,AN1,AN2
        mov     #0x0007,w1
        mov     w1,ADCSSL
        mov     #0x1FFF,w13
        bset    IPC2,#ADIP0
        bclr    IPC2,#ADIP1
        bclr    IPC2,#ADIP2
        bclr    IFS0,#ADIF
        clr     PORTE
        clr     PORTB
        mov     #0x07,w3
        mov     w3,TRISB
        clr     TRISE
        bclr    TRISD,#0
        clr     TRISF
        clr     PORTF
        clr     PORTD
        clr     PORTC
        bclr    TRISC,#13
        return
LCDSRT:
        call    BEK100
        mov.b   #0x03,w0
        call    YAZ
        call    ONAY
        call    BEK100
        mov.b   #0x03,w0
        call    YAZ
        call    ONAY
        call    BEK100
        mov.b   #0x03,W0
        call    YAZ
        call    ONAY
        mov.b   #0x02,W0
        call    YAZ
        call    ONAY
        mov.b   #0x02,W0
        call    YAZ
        call    ONAY
        mov.b   #0x08,W0
        call    YAZ
        call    ONAY
        mov.b   #0x00,W0
        call    YAZ
        call    ONAY
        mov.b   #0x06,W0
        call    YAZ
        call    ONAY
        mov.b   #0x00,W0
        call    YAZ
        call    ONAY
        mov.b   #0x0C,W0
        call    YAZ
        call    ONAY
        mov.b   #0x00,w0
        call    YAZ
        call    ONAY
        mov.b   #0x01,w0
        call    YAZ
        call    ONAY
        call    BEK100
        return
BEK100:
        do      #1000,BK100
        repeat  #0x3FF
        nop
BK100:  nop
        return
LHZR:   call    GOST
        swap.b  w11
        and.b   w11,#0x0F,w1
        add.b   #0x30,w1
        call    LCDYAZ
        swap.b  w11
        and.b   w11,#0x0F,w1
        add.b   #0x30,w1
        call    LCDYAZ
        return
BINBCD:
        clr     LW1
BCD1:   sl.b    SAYI1
        rlc.b   LW1
        mov     LW1,w0
        add.b   #0x03,w0
        btsc    w0,#3
        mov     w0,LW1
        mov     LW1,w0
        add.b   #0x30,w0
        btsc    w0,#7
        mov     w0,LW1
        dec.b   w9,w9
        bra     nz,BCD1
        sl.b    SAYI1
        rlc.b   LW1
        return
        .end
```

Ömerül Faruk Özgüven was born in Malatya, Turkey, in 1963. He obtained his Bachelor's degree in Electronics & Communications Engineering from the Yildiz Technical University, İstanbul, Turkey in 1985. He received the M.S. and the Ph.D. degree in institute of science from the Yildiz Technical University, İstanbul, Turkey; 1987 and 1996, respectively. He was appointed as Assistant Professor in 1994, in the Department of Electrical and Electronics Engineering, the Engineering Faculty of Inonu University, Malatya, Turkey. He is currently working as Assistant Professor in the Department of Biomedical Engineering, the Engineering Faculty of Inonu University, Malatya, Turkey. His research interests include fuzzy neural network, Digital Electronics, Microcontroller, Microprocessors, Embedded Systems, Programmable Logic Controller (PLC) and Industrial Applications.