



IMPLEMENTATION OF A PATH DETECTION ALGORITHM FOR MOVING IMAGES ON THE CNN UNIVERSAL MACHINE

Emel ARSLAN¹, Zeynep ORMAN², Sabri ARIK²

¹ Istanbul University, Research and Application Center for Computer Sciences, 34452, Beyazit, Istanbul, Turkey

² Istanbul University, Engineering Faculty, Computer Eng. Dept., 34320, Avcilar, Istanbul, Turkey
earslan@istanbul.edu.tr, ormanz@istanbul.edu.tr, ariks@istanbul.edu.tr

Abstract: CNN Universal Machines that contain two different processors working interactively with each other, have an important impact for image processing applications with their advanced computing features. These processors are named as ACE16k processor which is the hardware implementation of cellular neural networks and Digital Signal Processor (DSP). Bio-inspired (Bi-i) Cellular Vision System is also a CNN Universal Machine and its standalone architecture is built on CNN-type (ACE16k) and DSP-type microprocessors. In this study, certain objects in moving images are detected and their features are extracted. By using these features, an algorithm that finds out the path of moving objects is implemented on the Bi-i Cellular Vision System. Finally, the output images obtained as a result of this implementation are evaluated.

Keywords: CNN Universal Machine, Bi-i Cellular Vision System, ACE16k, Digital Signal Processor, Target Tracking.

1. Introduction

Image processing is one of the most important research topics in recent years. It is widely used in areas such as military, security, health, biology, astronomy, archeology and industry [1, 2]. For an image to be processed, it should first be presented in a format that a computer can understand, this means it should be converted into its related digital form. In the digital form, each of its pixel is expressed by means of the corresponding element of a matrix.

Algorithms that are developed for digital image processing require fast systems due to their processing load. Although conventional computer systems are in an increasing trend in terms of their speed, they are insufficient when image dimensions are getting larger, because they can process transactions in a serial manner. These computers cannot satisfy the need for speed, when we especially consider the implementation of real-time moving image processing algorithms that require at least 15-25 frames to be processed in a second.

Cellular Neural Network (CNN) theory that was proposed by Chua and Yang in 1988, is an analog, nonlinear and a real-time processing neural network model [3]. CNNs also have advanced features for image processing applications. In 1993, Roska and Chua have presented the CNN Universal Machine [4]. This analogical array computer has cellular processors (ACE4k, ACE16k, etc. [5, 6]) which are the hardware implementation of CNNs and it is very suitable for

image processing applications with its advanced computing capabilities. Bi-i Cellular Vision System is a CNN Universal Machine that can process high-speed and real time transactions and can be defined as a compact, independent and intelligent camera. This system has high-resolution sensors and two different processors named as CNN (ACE16k) and Digital Signal Processor (DSP) that can communicate with each other [7, 8].

In this study, firstly, an algorithm that finds out the path of certain objects in moving images is proposed. This algorithm initially recognizes certain objects in moving images and extracts their features. Then, it detects the path of these objects which they are tracking all along the image by using these extracted features. Secondly, an implementation of this algorithm is presented on the Bi-i Cellular Vision System and finally, the output images which are obtained as a result of this study are evaluated.

Moving target tracking is an important research field that is frequently used in robotics, video surveillance, traffic control, etc. Generally, there are two kinds of approaches adopted for target tracking [2].

- Recognition-based target tracking
- Motion-based target tracking

Recognition-based target tracking methods extract features and match them between image frames to fit the global motion model of moving image sequence. On the other hand, motion-based methods track image features over a sequence of images by using the difference between

each frame. Although the latter one has an approved theory foundation, the former one is more popular.

In recent years, there are numerous applications in the literature that use different methods for tracking moving objects. Kalman filtering is a widely used method for target tracking systems [2]. Some different approaches were also presented in [9] and [10] for traffic control and target tracking. [11] and [12] proposed different systems that detected moving objects on a mobile robot. In [13], Yilmaz et al. presented a detailed survey on object tracking and discussed and classified various related algorithms.

The remainder of this letter is organized as follows. Section 2 introduces fundamental concepts about CNN architecture, CNN Universal Machine, ACE16k processor, Bi-i Vision System and Bi-i programming, respectively. In Section 3, an algorithm that detects the path of certain objects in moving images is proposed and in Section 4, this algorithm is implemented on the Bi-i Vision System. Section 5 evaluates the obtained results and finally the paper is summarized in Section 6.

2. CNN Architecture and Bi-i Cellular Vision System

This section provides fundamental concepts about CNN architecture and Bi-i Cellular Vision System.

2.1. Architecture of the Cellular Neural Networks

Cellular Neural Networks (CNNs) derived from Hopfield Neural network is introduced in [3]. The two most fundamental components of the CNN paradigm are the use of analog processing cells with continuous signal values, and local interaction within a finite radius.

The structural design of CNN is formed with basic circuits called cells. Each cell containing a linear capacitor, a non-linear voltage controlled current source, and a few resistive linear circuit elements is connected to its neighboring cells; therefore direct interactions take place only among adjacent cells. Mathematical expression of the standard CNN model is described by the set of linear differential equations given by (1) which are associated with the cells in the circuit. The activation function of the CNN cell can be expressed by the set of nonlinear equation (2).

$$C \frac{d}{dt} x_{ij}(t) = -\frac{1}{R} x_{ij}(t) + \sum_{C(k,l) \in S_r(i,j)} A(i,j;k,l) y_{kl}(t) + \sum_{C(k,l) \in S_r(i,j)} B(i,j;k,l) u_{kl}(t) + z_{ij} \quad (1)$$

$$y_{ij} = f(x_{ij}) = \frac{1}{2} (|x_{ij} + 1| - |x_{ij} - 1|) \quad (2)$$

where,

$x_{ij} \in R$; State variable of cell $C(i,j)$,

$y_{kl} \in R$; Outputs of cells,

$u_{kl} \in R$; Inputs of cells,

$z_{ij} \in R$; Threshold,

$A(i,j;k,l)$; Feedback operator,

$B(i,j;k,l)$; Control operator.

y_{ij} ; Output equation.

Without loss of generality, linear resistor (R) and linear capacitor (C) values can be set to 1. The block diagram of a cell $C(i,j)$ is shown in the Figure 1.

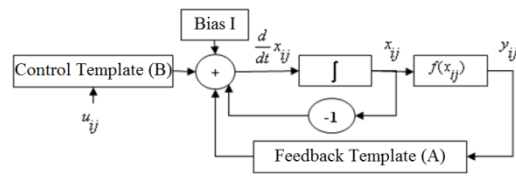


Figure 1. The block diagram of a cell $C(i, j)$

In this way CNNs have provided an ideal framework for programmable analog array computing. It means that the CNN can be used as a programmable device where the instructions are represented by the templates which define the connections between a cell and its neighboring cells. In general, the CNN templates consists of the feedback template (B), control template (A) and bias value. Basically, three different images can describe a CNN layer, that is, the input U, the state X and the output Y.

Let us assume a CNN with $M \times N$ cells and the cell in row i and column j is denoted as $C(i,j)$ [3]. r -neighborhood of a $C(i,j)$ cell is defined with the following definition (3) provided that r is a positive value.

$$N_r(i,j) = \left\{ C(k,l) \mid \max_{1 \leq k \leq M, 1 \leq l \leq N} \{ |k-i|, |l-j| \} \leq r \right\} \quad (3)$$

2.2. CNN Universal Machine

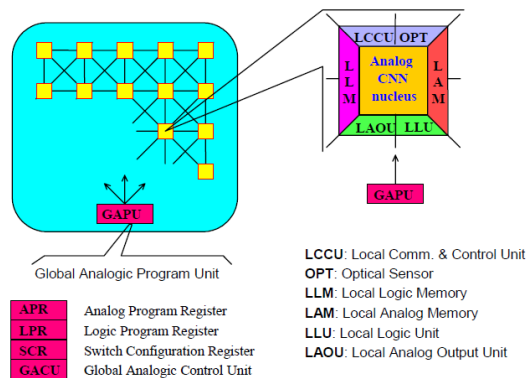


Figure 2. The architecture of the CNN-UM

The hardware implementation of CNN is easier compared to the Artificial Neural Networks as there is only connection between the neighbor cells and the cell structure. Analogical Cellular Engines (ACE4k, ACE16k etc. [6]) are based on CNN Universal Machine architecture. CNN Universal Machine(CNN-UM) architecture has been called by Roska and Chua as analogical computation since it can both perform the analog array operations and the logical operations together [4].

Figure 2 denotes the CNN-UM architecture which is based on the dynamic computing of a simple CNN. This figure shows the elements in the complex CNN Nucleus and the functional blocks of the Global Analogic Programming Unit[8].

2.3. ACE16k Processor

ACE16k, is a CNN based processor of CNN Universal Machine which can perform analog operations. ACE16k which is used to perform various image processing operations contain low resolution (128 x 128) CMOS gray level image sensor and analog processor arrays. This processor array is much faster (30000 frames per second) than the conventional processors in image processing applications since it can processes the whole image in parallel.

2.4. Bi-i Cellular Vision System

The Bi-i Cellular Vision System which contains two different processors, a CNN based ACE16k and a DSP that can be defined as a compact, standalone and intelligent camera capable of real time operations at very high speed [7, 8]. The images are stored in local memories with the help of two different sensors as a (1280x1024) color CMOS sensor, and a (128x128) ACE16K sensor.

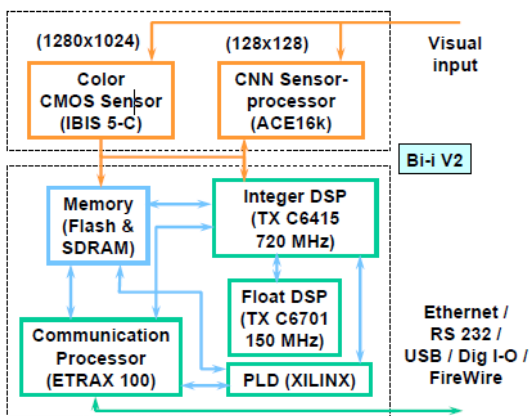


Figure 3. The block diagram of the Bi-i V2 Cellular Vision System

The block diagram of the Bi-i V2 Cellular Vision System is given in Figure 3. As seen from the figure, this system has a color CMOS sensor array (IBIS 5-C) and two high-end digital signal processors (TX C6415 and TX C6701). This system runs an embedded Linux

on the communication processor and has complex external interfaces like USB, FireWire and a general digital I/O in addition to the Ethernet and RS232[8].

2.5. Bi-i Programming

CNN Universal Machine has two different programming methods. One of them is AMC (Analogical Macro Code) language which is a conventional Bi-i programming method. The codes written in AMC language are converted to binary basis and run on Bi-i. Another method is the Bi-i (Software Development Kit - SDK) which is used to develop more complex applications. Bi-i SDK, consists of the C++ programming library which is a group used to develop applications. These libraries can also used for the Digital Signal Processor (DSP) with the development unit Code Composer Studio and they contain many functions to control the whole ACE16k circuit [7].

3. An Algorithm that Detects the Path of Objects in Moving Images

In this section, we will present an algorithm that detects the path of shifted objects in moving images. The moving image we consider is a traffic flow taken on a highway. This algorithm has been successfully implemented on traffic scenes and it has detected the path of certain objects such as cars, trucks, etc. The colored moving image we consider, is first preprocessed so that it is transformed in a format that the functions can process to detect certain objects and the path of each object in the image. We will discuss these preprocessing operations in the Preprocessing Section in detail.

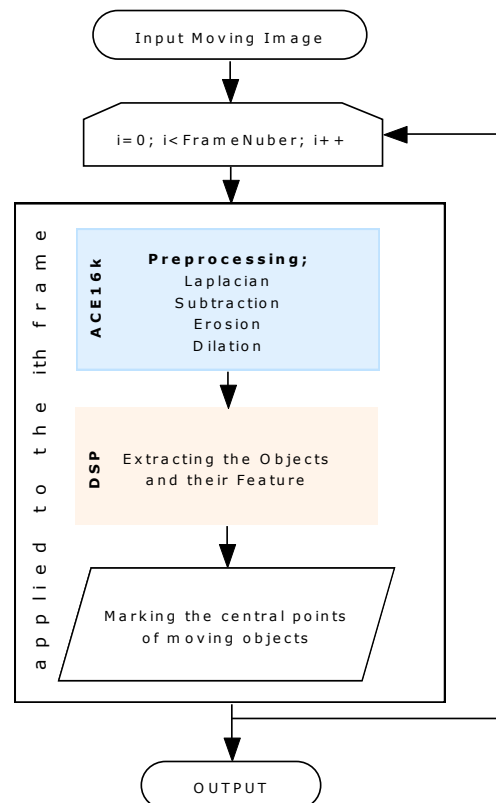


Figure 4. An algorithm that detects the path of moving objects

After preprocessing phase, the moving image can be an input to a function that finds out the objects and extracts their features from the image. This function produces an output of extracted features of the detected objects. These features will then be an input to another function that detects moving objects in the image.

Then, using these extracted features, locations of each moving object are marked in each frame. When this process is complete for all frames, the path that a certain moving object is tracking all along the image will be plotted.

Figure 4 denotes the steps of the algorithm that is used for this implementation. As clearly seen from the algorithm, preprocessing phase is run on the ACE16k processor and other operations run on the DSP processor. Therefore, the algorithm is implemented on both processors interactively for different phases which makes the processing of images faster.

4. Implementation of the Algorithm on the Bi-i Vision System

In this section, the algorithm that detects the path of shifted objects in moving images will be implemented on the Bi-i Cellular Vision System. The phases of this implementation is given in detail as follows:

4.1. Preprocessing Phase

The moving image we consider for our implementation, first undergoes a series of enhancement operations to get better results in detection of objects.

Because the moving image is colored, we first transform it into a gray level image by using RGB2ByteMatrix function and then apply the preprocessing operations. To increase the success of detecting objects in a moving image, we use a Laplacian based image enhancement technique so that the edges in the image have been made more apparent [2].

Image enhancement technique based on Laplacian can be implemented by using a Laplace template or a Laplace function that is defined in Bi-i SDK [7]. Laplacian is a second order differential operator and it enhances gray level discontinuous points of an image, this means the edges of the image become much clearer by using this technique. It also reduces the importance of slowly changing gray level areas. The resulting image after we obtain by applying the Laplace function, is subtracted from the base image pixel by pixel. In this manner, the image is enhanced to get better edges in the next step.

The Laplace function and the Subtract function that is used to subtract the resulting image from the original input are both ran on the ACE16k processor.

In Figure 5, original images and gray level images that is obtained during the preprocessing phase of the implementation for different frames are given.

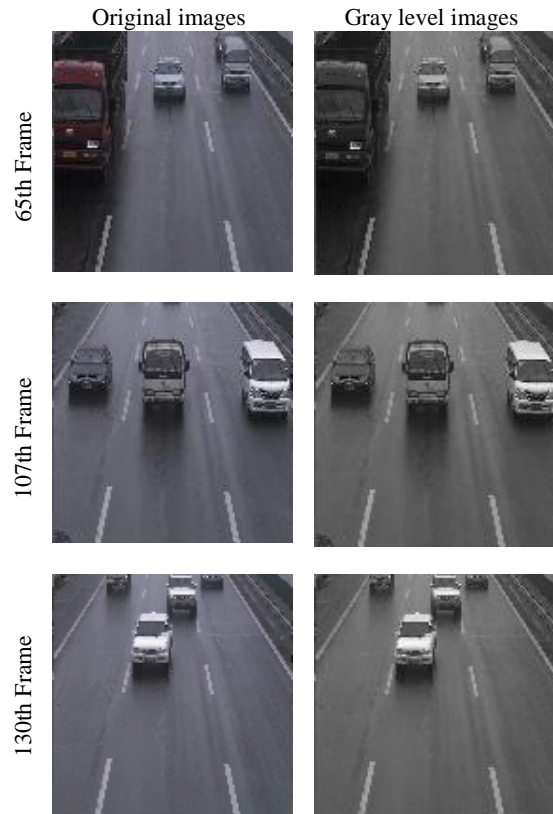


Figure 5. Original images and gray level images for different frames

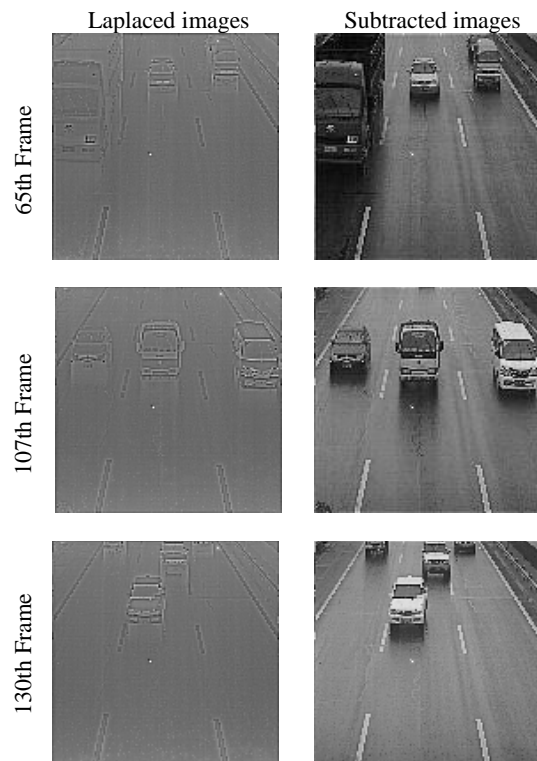


Figure 6. The resulting images that are obtained after the Laplacian and Subtraction phases

Figure 6 shows the Laplacian and the subtracted images of the original images which are obtained during the implementation of the Laplacian-based image enhancement procedures for different frames.

After the enhanced moving image is converted into a binary image with the ByteMatrix2BitMatrix, it undergoes a 3 iteration erosion and a 2 iteration dilation, respectively. With these operations, the objects become more apparent. Figure 7 shows the resulting images that are obtained after the erosion and dilation phases.

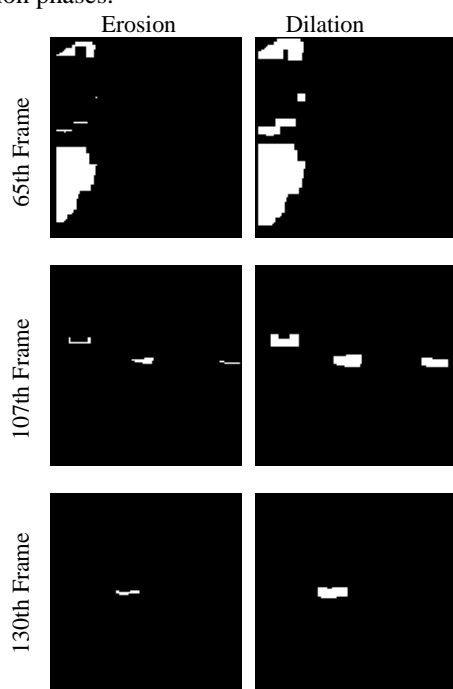


Figure 7. The resulting images that are obtained after the Erosion and Dilation phases

After the preprocessing phase is completed, the moving image is now ready for the CalcFeatures function to be an input image which will detect the objects and extract the features of these objects.

4.2. Extracting the Objects and their Features

The CalcFeatures function in the Instant Vision library is used to detect the objects in each frame and extract the features of these objects from the preprocessed binary image. This function can output the features like area, boundary values, extremes, orientation, central point, eccentricity, diameter, etc. that belong to each object in the image.

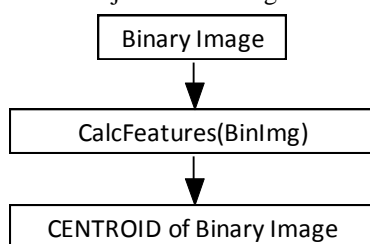


Figure 8. Feature extraction scheme

FEAT_CENTROID parameter of the CalcFeatures function indicates that the this function should only calculate the central point of the objects. It is very important in our implementation to know the central point of the objects in each frame to find out the path that these objects are tracking. We've just used FEAT_CENTROID Feature Selection parameter because there is no need to calculate the other features. Figure 8 shows the basic steps of the feature extraction phase.

The objects which their central points are calculated are now ready for the operations that will detect the path of these objects. The functions that can perform these operations are defined in MTT (Multi Target Tracking) which is located under the Instant Vision library.

Meas vector variable which contains the central point information of an object is given as an input parameter of a function that will track the moving objects in the image. By using this vector, this function also checks all the objects that are in the path tracer of the Tracker function. If a specif object is detected as a moving object, the function that is written in C++ programming language, will mark the central point of this object. After this operation is performed for every object in each frame, the path of each object will be plotted in the moving image.

5. Experimental Results

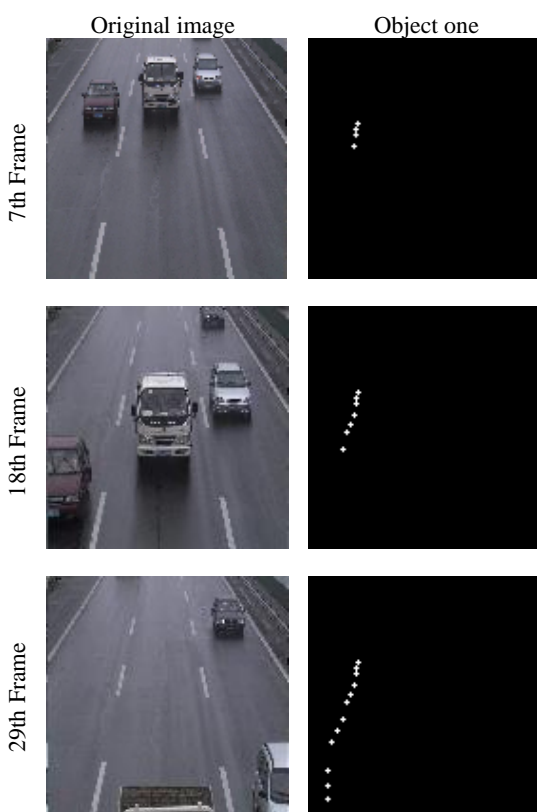


Figure 9. Sample output images showing the path of the object one

For this implementation, the moving image we consider is a traffic flow taken on a highway. Therefore, each moving object in the image corresponds to a vehicle such as cars, trucks, etc. When the central point of the moving

objects in each frame is marked, it will give us the path that they are tracking all along the highway.

When the algorithm which is detailed in Section 3, is implemented on the Bi-i Vision System, some sample output images can be obtained as given in Figure 9 and 10 for object one and object two, respectively.

In the original images that are given in Figure 9 and Figure 10, the red car which is going along the leftmost traffic lane is named as the object one and the truck which is going along the central traffic lane is named as object two. In each frame, the central points of all objects are marked. As clearly seen from both figures, the central points of object one and object two are marked for the 7th, 18th and 29th frames. After this operation is complete for all frames that contain the moving object, we obtain the path plotted in the image.



Figure 10. Sample output images showing the path of the object two

6. Conclusion

In this paper, we have presented an algorithm that detects certain objects and extracts their features in moving images and with these features, this algorithm finds out the path of the moving objects. We have also implemented this algorithm by using the Bi-i Cellular Vision System which offers an ultra high speed processing in image processing applications. The preprocessing operations of the algorithm which are Laplacian image enhancement, subtraction, dilation and erosion are implemented on the ACE16k processor, respectively. The other operations that are

converting the image into a gray level image, extracting features of the objects, detecting moving objects by using the extracted features and detecting the path of these objects in the image are all implemented on the Digital Signal Processor (DSP). When tested in traffic scenes, the algorithm successfully detects the path of each moving object in the moving image by generating different image files.

This study shows that the Bi-i Vision System which is successfully used in image processing applications, can also give very fruitful results for moving and multi-target tracking applications.

For a future work, the algorithm that we proposed for path detection can be improved to determine the traffic faults like lane violation of the vehicles in a traffic flow on a highway.

7. References

- [1] C. Gonzales and R.E. Woods, "Digital Image Processing", Prentice Hall, New Jersey, 2002.
- [2] T. Acharya and A.K. Ray, "Image Processing: Principles and Applications", Wiley and Sons, 2005.
- [3] L. O. Chua and L. Yang, "Cellular neural networks: Theory and applications", IEEE Trans. on CAS, vol. 35 no. 10, pp.1257–1290, 1988.
- [4] T. Roska and L. O. Chua, "The CNN universal machine: an analogic array computer", IEEE Trans. on CAS-I, vol. 40 no.3, pp. 163–173, 1993.
- [5] T. Roska and A. Rodriguez-Vazquez, "Towards visual microprocessors", Proceedings of the IEEE, vol. 90 no.7, pp. 1244–1257, 2002.
- [6] A.R. Vazquez, G. L. Cembrano, L. Carranza, E.R. Moreno, R.C. Galan, F.J. Garrido, R.D. Castro and S. E. Meana, "ACE16k: the third generation of mixed-signal SIMDCNN ACE chips toward VSoCs", IEEE Trans. CAS-I, vol. 51, no.5, pp. 851–863, 2004.
- [7] <http://www.analogic-computers.com/Support>
- [8] A. Zarandy and C. Rekeczky, "Bi-i: a standalone ultra high speed cellular vision system.", IEEE Circuit and Systems Magazine, vol. 5, no.2, pp. 36–45, 2005.
- [9] CUCCHIARA, R., GRANA, C., NERI, G., PICCARDI, M., and PRATI, A., The Sakbot System for Moving Object Detection and Tracking, Video-Based Surveillance Systems- Computer Vision and Distributed Processing, pp.145-157, 2001.
- [10] ROTA, N., and THONNAT, M., Video sequence interpretation for visual surveillance, Proc. of Third IEEE International Workshop on Visual Surveillance, pp. 59-68, 2000.
- [11] JUNG, B. and SUKHATME, G.S., Detecting Moving Objects using a Single Camera on a Mobile Robot in an Outdoor Environment, In the 8th Conference on Intelligent Autonomous Systems, pp. 980-987, Amsterdam, The Netherlands, March 10-13, 2004.
- [12] BEHRAD, A., SHAHROKNI, A. and MOTAMEDI, S.A., A robust vision-based moving target detection and tracking system. In the Proceeding of Image and Vision Computing Conference, University of Otago, Dunedin, New Zealand, November, 2001.
- [13] YILMAZ, A., JAVED, O., and SHAH, M., Object tracking: A survey. ACM Comput. Surv. 38(4), Article 13, 45 pages, 2006.

Note:

Emel Arslan 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 a Computer Engineer in the Research and Application Center for Computer Sciences, Istanbul University. Her research interests are artificial neural networks, natural language processing, image processing applications and intelligent systems.



Zeynep Orman received the B.Sc., M.Sc. and Ph.D. degrees from Istanbul University, Istanbul, Turkey, in 2001, 2003 and 2007, respectively. She has studied as a postdoctoral research fellow in the Department of Information Systems and Computing, Brunel University, London, UK in 2009. She is currently working as an Assistant Professor in the Department of Computer Engineering, Istanbul University. Her research interests are artificial neural networks, nonlinear systems, image processing applications and intelligent systems.



Sabri Arik received the Dipl.Ing. degree from Istanbul Technical University, Istanbul, Turkey, the Ph.D. degree from the London South Bank University, London, UK, and the Habilitation degree from Istanbul University, Istanbul, Turkey. He is now with the Department of Computer Engineering, Istanbul University as a Professor. His major research interests include cellular neural networks, nonlinear systems and matrix theory. He has authored and coauthored some 50 publications. Dr. Arik is a member of the IEEE Circuits and Systems Society Technical Committee of Cellular Neural Networks and Array Computing. He was the recipient of the Outstanding Young Scientist Award in 2002 from the Turkish Academy of Sciences, Junior Science Award in 2005 from the Scientific and Technological Research Council of Turkey and the Frank May Prize (Best Paper Award) in 1996 from the London South Bank University.

