Gazi University Journal of Science
Part A: Engineering And Innovation
GU J Sci Part:A
3(1):15-25 (2015)

Object Recognition with SIFT and MI-SIFT Methods

Fırat HARDALAÇ[1], Abdullah ORMAN[2,♠], Berkan URAL[3], Ali EREN[4]

[1]*Gazi University, Electrical Electronics Engineering, Ankara, TURKEY*
[2,*]*Gazi University, Graduate School of Natural and Applied Sciences, Ankara, TURKEY*
[3]*Gazi University, Electrical Electronics Engineering, Ankara, TURKEY*
[4]*Kastamonu University, Information Technology Department, Kastamanu, TURKEY*

Received: 15/01/2015 Accepted: 06/04/2015

ABSTRACT

This study, which has been commonly used in object recognition area for many years, is based on the recognition of special point-based objects. There are a lot of differences in some parameters such as luminance, anguler and resistance factors from cyclic and dimensional changes. Generally, independent feature transformation (mirror reflection invariant feature- MIFT) method which is a weak point version of SIFT method is created through the development process against the mirroring effect and scalar invariant feature transformation (scale invariant feature transform- SIFT) method is investigated in this experiment in detail. In this work, 11 sample trials with images which are in different cases are performed one by one respectively to analyze these methods in detail and this gives us an opportunity to compare the results of these methods clearly. The aim of this work is, understanding the performance and the capability of the object recognition methods, which are SIFT and MIFT, in time domain and space domain areas, also determining the speed value and the acqusition time parameter of these methods represents us to observe the advantages and disadvantages of SIFT and MIFT methods.

Key Words: *Object recognition, keypoint descriptors, scale invariant feature transform, SIFT, mirror reflection invariant feature transform, MIFT.*

1. INTRODUCTION

Object recognition which is located in the center of the natural learning process, has been gaining importance abruptly and also this can be used in many areas today. With the development of the technology, this innovation has been integrated into the some areas and this has been the basis of many devices such as optical cameras, security cameras [1,2].

Several methods have been used in object recognition, also most of these have been commonly used in "background elicitation", "partitioning" and "specific point identifiers". Among these methods, SIFT and MIFT which are used in scientific works are the basis of this study.

In this work, scale invariant feature transform(SIFT) method which is in the group of special-based identifiers is analyzed in detail.Background elicitation method is widely used for detecting a moving object. For the method of "Defining a special point", there are lots of methods such as FIND, SPIN, SURF, GLOH, DAISY. Among these methods, better results are obtained from SIFT method, also these observed data are

♠Corresponding author, e-mail: abdullah@gazi.edu.tr

least affected from some parameters such as brightness, dimensional and cyclical transformations. Defining a special point method is generally consisted from 3 main stages. These stages are shown as a diagram in Figure 1 below.

Figure 1: Stages of defining a special point process

In the first step of the process, edges and other regions are identified and image analysis is done in general. For the analyzing process, generally, some complex structures are used such as Harris Point, Harris-Laplace regions, Hessian-Laplace regions, Difference of Gaussian (DoG) [3,4]. In the second part of the process, regions are classified and vector and spatial data are analyzed. In this step, process is done by using SIFT, GLOH, PCA-SIFT, shape context, compex filters and steerable filters [5,6].

In the last part of the process, points which have the same coordinates, are used and location equivalent points are obtained. Then, comparisons are done and mutual pairing process is achieved. For this work, SIFT method is used as an identifier and this is developed with using the distribution of an image's data.

2. MATERIALS AND METHODS

In this study, two different applications and several comparisons were made. For the first application, six different images from an image which was one glasses were used. For the second application, five different images from an image which was a table and a pen object were used. For each image, using the SIFT and MIFT methods, object modeling process was performed under different conditions and comparisons were made by using the obtained results. In addition, these simulations were done on a Macbook Air which has i5 CPU and OSX Yosemite Software.

2.1. SIFT Method

SIFT method which is based on the gradient information of an image is one of the important methods that is used in object recognition with a particular point. This method shows resistivity to some factors such as dimensional, circular, brightness and three-dimensional changes [7,8].

In this method, firstly, in order to create a smooth pyramid array, the image and a Gaussian kernel are inserted in the time domain and the convolution provides the image to resample itself again [9,10]. As a result, the output image is applied in the space scale of the pyramid of Difference of Gaussian (DoG), also other key points are determined by using the maximum and minimum values [11,12]. After these processes, SIFT descriptor can be independent from the scale parameter [13]. For observing the dominant orientation of each key point, gradient histograms of the neighborhood in the image part are evaluated individually. Also, key points can have scale, position and orientation information.

2.1.1 Algorithm

SIFT method is analyzed in 7 steps generally.

1- Determining the scalar space and limits.
2- Determining the minimum and maximum values by using Difference of Gaussian(DoG).
3- Determining the end points.
4- Locating key points in the pixels.
5- Normalization process with using filters.
6- Gaining resistance to cyclical changes.
7- Finding the key point descriptors.

2.1.1.1 Determining the Scalar Space

The scalar space of an image is a function which is equal to the convolution with different sized Gaussian kernel functions of an image[14]. The expression (1) is given below.

$$L(x, y, k\sigma) = G(x, y, k\sigma) * I(x, y) \qquad (1)$$

Gauss kernel function is associated with the expression (2) which is given below.

$$G(x, y, k\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x2+y2)}{2\sigma^2}} \qquad (2)$$

In SIFT algorithm, when determining the scalar space, Difference of Gaussian(DoG) method is generally used. In this method, the original image is filtered by a Gauss filter which has a different standard deviation value [15]. After this step, Gauss pyramid is obtained. Then, the image is subtracted from the other obtained images and at the end of this stage DoG images are seen clearly.

2.1.1.2 Determining the minimum and maximum values by using Difference of Gaussian(DoG).

Key points which has the scale and location information are important parts of the image. For observing the dominant orientation of the key points, gradient histograms of the neighborhood are drawn individually[16]. When disposing the key points, end points and calculated locations are interpolated and this gives a more stable structure[17].

The second order Taylor series of the DoG function is given below(3).

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \qquad (3)$$

For this equation, end points are calculated by the derivative of the D(x) function and this function must equal to zero. This equation is given below (4)(5).

$$x_1 = -\frac{\partial^2 D^{-1}}{\partial x^2}\frac{\partial D}{\partial x} \quad (4)$$

$$D(x_1) = D + \frac{1}{2}\frac{\partial D^T}{\partial x}x_1 \quad (5)$$

2.1.1.3 Gaining Resistance to Cyclical Changes and Finding the Key Point Identifiers

An area is chosen according to the calculated data and after this step, smoothing image(L), gradient value (m(x,y)) and angle value (Θ(x,y)) are obtained. For calculating the gradient value and the angle value, the expressions (6),(7) which are given below are generally used.

$$m(x,y) = \sqrt{(L(x+1,y)-L(x-1,y))^2 + (L(x,y+1)-L(x,y-1))^2} \quad (6)$$

$$\theta(x,y) = \tan^{-1}\frac{(L(x+1,y)-L(x-1),y))}{(L(x,y+1)-L(x,y-1))} \quad (7)$$

2.2. MIFT Method

MIFT (Mirror Reflection Invariant Feature Transform) method is used to get better results than SIFT method in mirroring process[18]. The mirroring method which is used in this study is performed in three different ways. These are; the horizontal mirroring, the vertical mirroring, the complex mirroring. SIFT method is independent from the rotation of the coordinate system, so this method gives the same results for both the horizontal and vertical mirroring [19].

2.2.1 Algorithm

When the original image and the horizontal mirrored image are analyzed, it is observed that the pixels are arranged in an order and column arrangement of the pixels is reversed. In Figure-2, some cases are given respectively, a key point which is identified on an image(a), observing the new pixel arrangement for the case of mirroring(b), the orientation of key points in the original image(c), the orientation of key points in mirrored image(d), the orientation of key points for using SIFT method(e), the orientation of key points in mirrored image for using SIFT method(f), the orientation of key points in mirrored image for using MIFT method(g)[20].

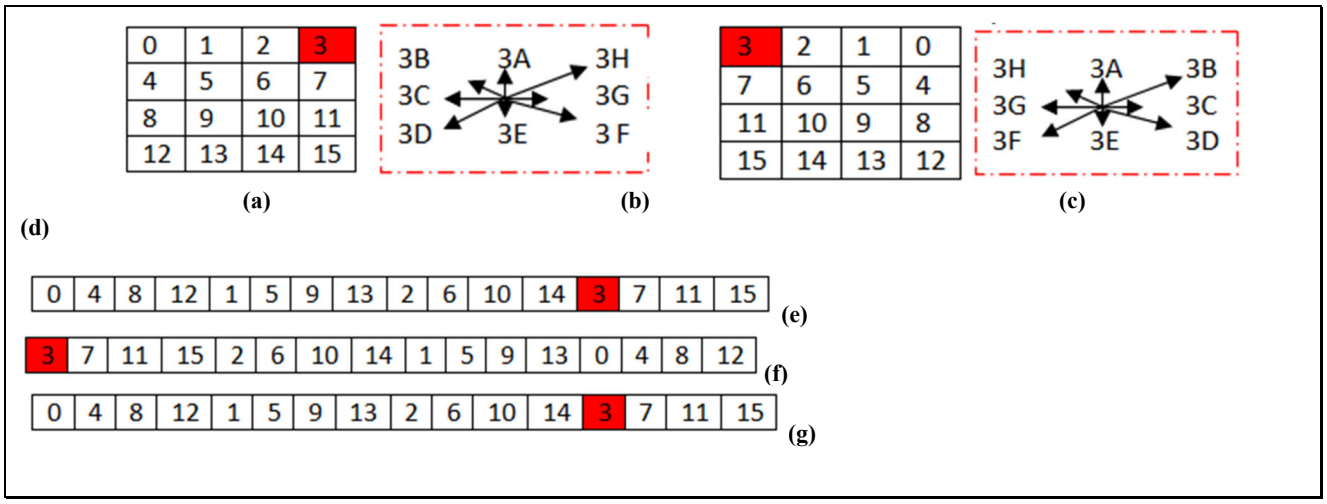
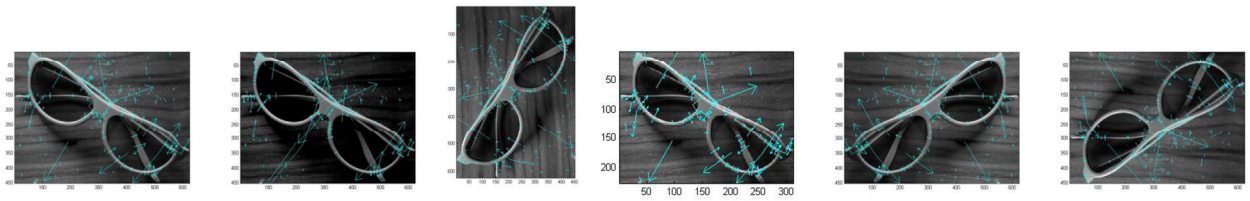


Figure 2: A key point which is identified on an image(a), observing the new pixel arrangement for the case of mirroring(b), the orientation of key points in the original image(c), the orientation of key points in mirrored image(d), the orientation of key points for using SIFT method(e), the orientation of key points in mirrored image for using SIFT method(f), the orientation of key points in mirrored image for using MIFT method(g)

3. RESULTS AND DISCUSSION

In this study, two applications were carried out to obtain the image data. In this study; 6 different images were used in the first part and 5 different images were used in the second part of the work. The applications and stages are given below in detail.

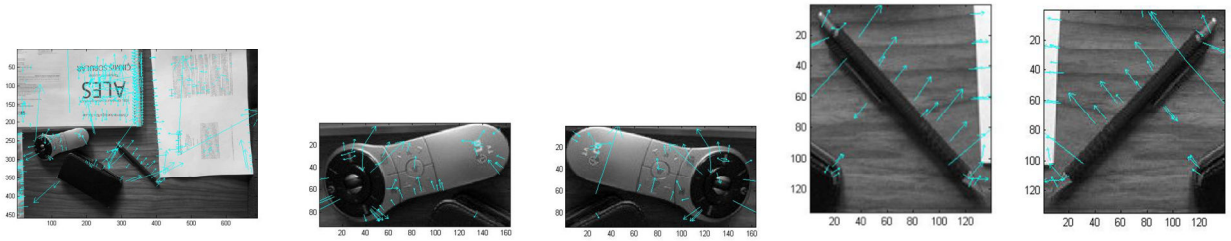
Experiment 1



(a) (b) (c) (d) (e) (f)

Figure-3:SIFT key points of and a) original image, b) a reduced gloss image, c) 90 degree rotated image, d) the reduced image, e) horizontal mirrorred image and f) vertical mirrorred images.

Experiment 2



(a) (b) (c) (d) (e)

Figure-4: SIFT key points representation of a) general view, b) the image part, c) result image of the mirrored part, d) the image of a different portion, e) mirrored image with using the new portion.

For the first experiment, matching was investigated between different cases of an object. For this regard, Figure 3 was consisted from a) original image, b) a reduced gloss image, c) 90 degree rotated image, d) the reduced image, e) horizontal mirrorred image and f) vertical mirrorred images, respectively.

For the second experiment, a complete picture was tried to be compared with its part. According to this operation, Figure 4 was given respectively, a) general view, b) the image part, c) result image of the mirrored part, d) the image of a different portion, e) mirrored image with using the new portion.

In these applications, firstly, key points and directions were obtained with using the first image in a different contrast condition. For this determination process, gradient histograms of the neighborhood which were in the image part were removed individually. Key point scale, location, orientation information were also shown clearly. Determining the key points in the regions was more difficult where contrast increased. Also, the point of orientation and scale parameters were dependent by the changes of the contrast value.

Then, SIFT method was performed on these images. Matching key points, their rotation and other results were displayed in Figure 5, Figure 6 and Table 1, respectively.

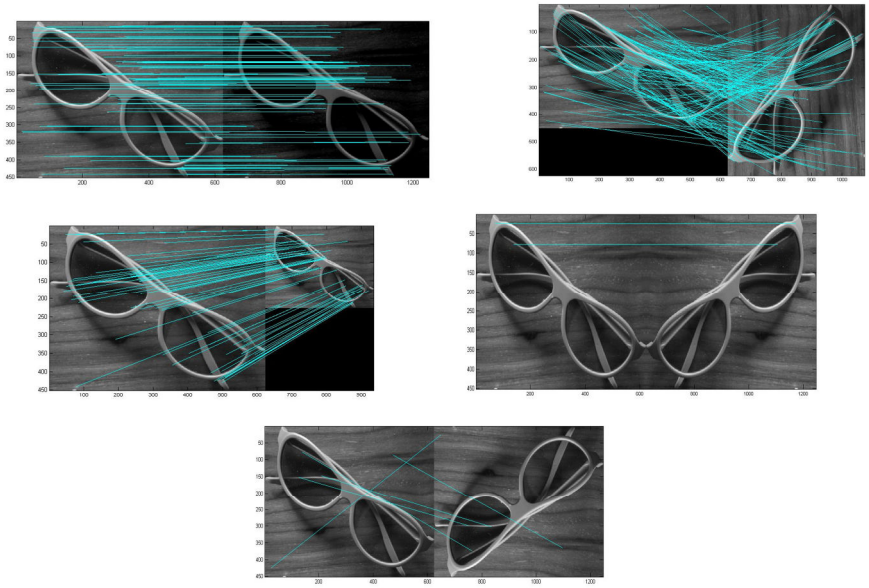


Figure-5: Mappings for Application-1 made by the SIFT method. a)a-b, b)a-c, c)a-d, d)a-e, e)a-f

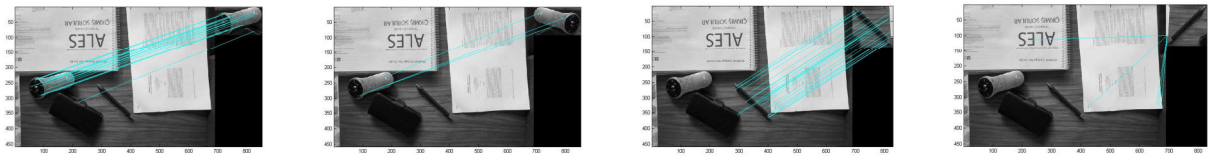


Figure-6: Mappings for Application-2 made by the SIFT method. a)a-b, b)a-c, c)a-d, d)a-e

Table-1: Matching results with using SIFT method

Matching Conditions for Application-1	**Matching Results for SIFT method**
Original image(a)-brightness modified image(b)	115
Original image(a)-rotated image(c)	194
Original image(a)-compressed image(d)	65
Original image(a)-horizontal mirroed image(e)	2
Original image(a)-vertical mirroed image(f)	6
Matching Conditions for Application-2	**Matching Results for SIFT method**
Original image(a)- brightness modified image(b)	21
Original image(a)-mirrored image of a part(c)	4
Original image(a)-image of different object(d)	26
Original image(a)-mirrored image of the object (e)	5

Table 1 shows that when the matching data value is higher; the object is accepted as well-known. Then, for the first image series (Application 1); SIFT and MIFT methods were applied and results were compared. Comparisons were made with tables and in these tables, “processing time” refers to scanning the relevant picture, “matching time” refers to points matching time, “point number” shows key point numbers of pictures and finally “matching ratio” refers to the result of dividing matching point to matching process time.

Case-1: general view (a) vs less bright image (b)

SIFT method image MIFT method image

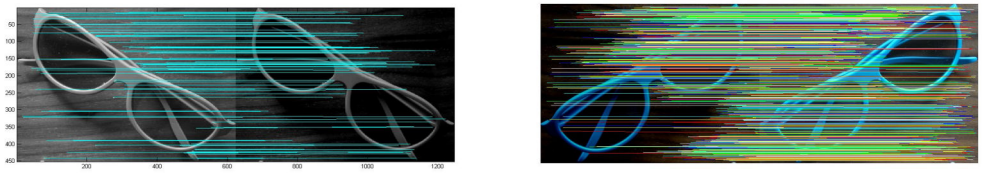


Table-2: Matching results of SIFT and MIFT methods for the general view (a) in 1th image series and the brigthness reduced image(b)

	SIFT method	MIFT method
1st picture processing time	1.0732 s	2.8418 s
2nd picture processing time	1.0690 s	2.9761 s
Matching process time	2.5398 s	2.4743 s
1st picture point number	239 nk	2485 nk
2nd picture point number	202 nk	2978 nk
Matching point number	115 nk	1396 nk
Matching ratio	0.26	0.255
Processing time per key points*100	0.575 s	0.045 s

Case-2: general view (a) vs rotated image (c)

SIFT method image MIFT method image

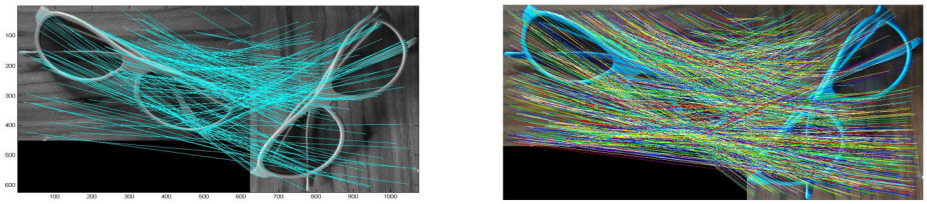


Table-3: Matching results of SIFT and MIFT methods for the general view (a) the 90 degree rotated image(c)

	SIFT method	MIFT method
1st picture processing time	1.0501 s	2.7826 s
2nd picture processing time	1.2082 s	2.7873 s
Matching process time	2.5631 s	2.1807 s
1st picture point number	239 nk	2485 nk
2nd picture point number	256 nk	2471 nk
Matching point number	194 nk	1607 nk
Matching ratio	0.39	0.325
Processing time per key points*100	0.517 s	0.044 s

Case-3: general view (a) vs compressed image (d)

SIFT method image MIFT method image

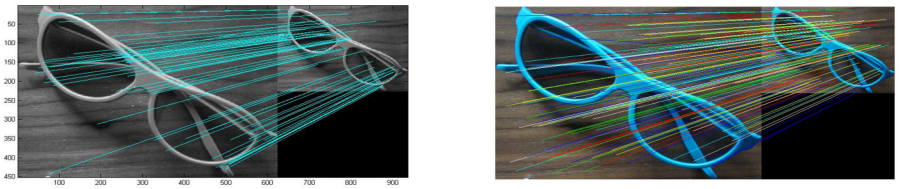


Table-4: Matching results of SIFT and MIFT methods for the general view (a) the compressed image(d)

	SIFT method	MIFT method
1st picture processing time	1,0602s	3.0141s
2nd picture processing time	0,3665s	0.6578s
Matching process time	1,7613s	0.8175s
1st picture point number	239nk	2485nk
2nd picture point number	103nk	503nk
Matching point number	65nk	207nk
Matching ratio	0,19	0,070
Processing time per key points*100	0,515 s	0,027 s

Case-4: general view (a) vs vertically mirrored image (e)

SIFT method image MIFT method image

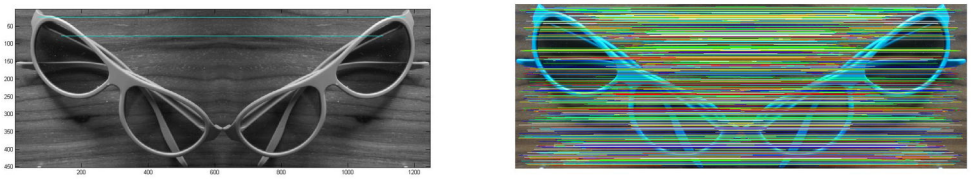


Table-5: Matching results of SIFT and MIFT methods for the general view(a) the horizontal mirrored image(e)

	SIFT method	**MIFT method**
1st picture processing time	1,0602s	3.0141s
2nd picture processing time	0,3665s	0.6578s
Matching process time	1,7613s	0.8175s
1st picture point number	239nk	2485nk
2nd picture point number	103nk	503nk
Matching point number	65nk	207nk
Matching ratio	0,19	0,070
Processing time per key points*100	0,515 s	0,027 s

Case-5: general view (a) vs horizontally mirrored image (f)

SIFT method image MIFT method image

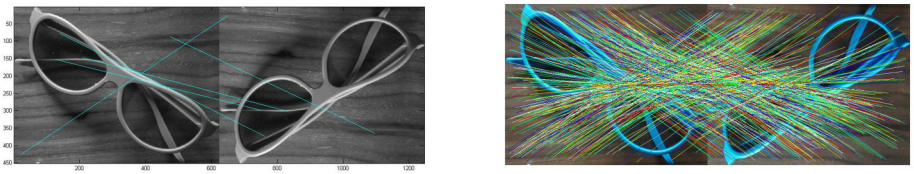


Table-6: Matching results of SIFT and MIFT methods for general view(a) vertical mirrored image(f)

	SIFT method	MIFT method
1st picture processing time	1.0816 s	2.7315 s
2nd picture processing time	1.2783 s	2.7282 s
Matching process time	2,5063 s	2,2198 s
1st picture point number	239nk	2485nk
2nd picture point number	241nk	2535nk
Matching point number	6nk	1122nk
Matching ratio	0,0125	0,224
Processing time per key points*100	0,522 s	0,044 s

At this stage, analyzing the matching results for the second image sequence(App-2) is the next step.

Case-1: general view (a) less bright part (b)

SIFT method image MIFT method image

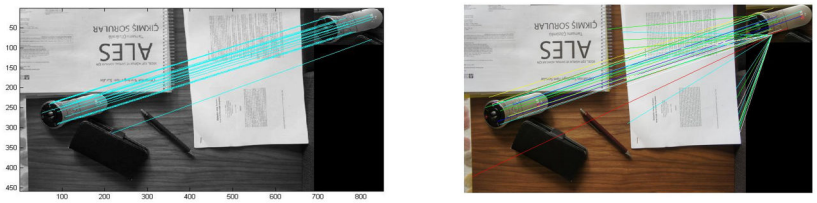


Table-7: Matching results of SIFT and MIFT methods for the general view(a2) the portion of the image(b2)

	SIFT method	**MIFT method**
1st picture processing time	1.5367 s	2.9225 s
2nd picture processing time	0.1619 s	0.1378 s
Matching process time	1.8350 s	0.4874 s
1st picture point number	709 nk	2063 nk
2nd picture point number	46 nk	232 nk
Matching point number	21 nk	118 nk
Matching ratio	0,0028	0,035

Case-2: general view (a) mirrored image (c)

SIFT method image MIFT method image

Table-8: Matching results of SIFT and MIFT methods for the general view(a1) the mirrored image(c2)

	SIFT method	**MIFT method**
1st picture processing time	1,5975 s	2.8636s
2nd picture processing time	0,1386 s	0.1222s
Matching process time	1,7483 s	0.4779s
1st picture point number	709 nk	2063nk
2nd picture point number	34 nk	72 nk
Matching point number	4 nk	14nk
Matching ratio	0,005	0,006
Processing time per key points*100	0,233 s	0,022 s

Case-3: general view (a) a different object from the image (d)

SIFT method image MIFT method image

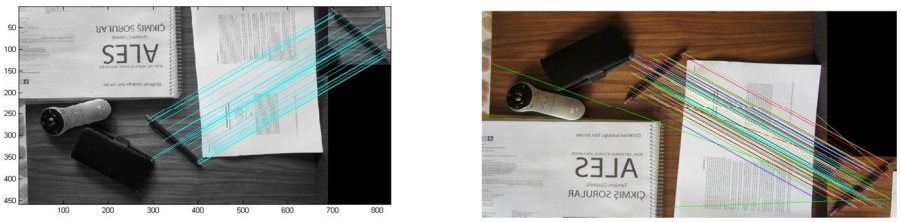


Table-9: Matching results of SIFT and MIFT methods for the general view(a1) the new portion of the image(d2)

	SIFT method	MIFT method
1st picture processing time	1,5377 s	2.8382s
2nd picture processing time	0,1279 s	0.2205s
Matching process time	1,7227 s	0.5547s
1st picture point number	709 nk	2063nk
2nd picture point number	36 nk	220 nk
Matching point number	5 nk	61nk
Matching ratio	0,006	0,027
Processing time per key points*100	0,231s	0,024s

Case-4: general view (a) mirrored image (e)

SIFT method image MIFT method image

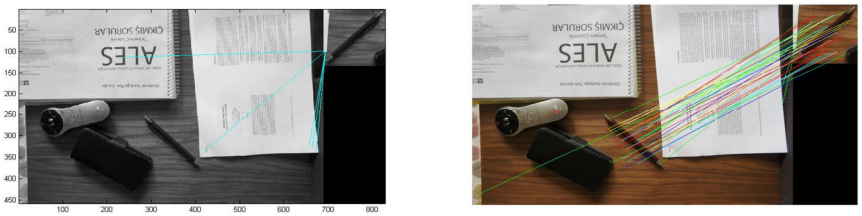


Table-10: Matching results of SIFT and MIFT methods for the general view(a1) the new mirrored object in the image(e2)

	SIFT method	MIFT method
1st picture processing time	1,5377 s	2.8382s
2nd picture processing time	0,1279 s	0.2205s
Matching process time	1,7227 s	0.5547s
1st picture point number	709 nk	2063nk
2nd picture point number	36 nk	220 nk
Matching point number	5 nk	61nk
Matching ratio	0,006	0,027

For the 1th and 2nd pictures, when we investigate the matching process time and rate and average time parameters; it is clearly seen that if we apply SIFT method to the original image and vertical mirrored image, matching becomes faster, because less number of points are matched. Also, if we apply this method again to the original image and a portion of the image, matching needs much more time, because there occurs a maximum capability of matching. Also, to define an object recognition in detail, MIFT method gives the best results and its matching rates are higher than the SIFT method; but for detecting, this method is required much more time than the other.

4. CONCLUSIONS AND FUTURE WORKS

After the matching process, 5 matching images were obtained from the 1st image sequence (App-1) and 4 matching images were obtained from the 2nd image sequence (App-2). Analyzing these sequences in detail, all results were presented in a statistical graphic format in Figure 7 and Figure 8.

In Figure 7; For first image with using SIFT method and for second image with using SIFT method; number of key points and matching points were given in this bar chart. Also, for the first and second images with using MIFT method, number of key points and matching points are given in this bar graph again.

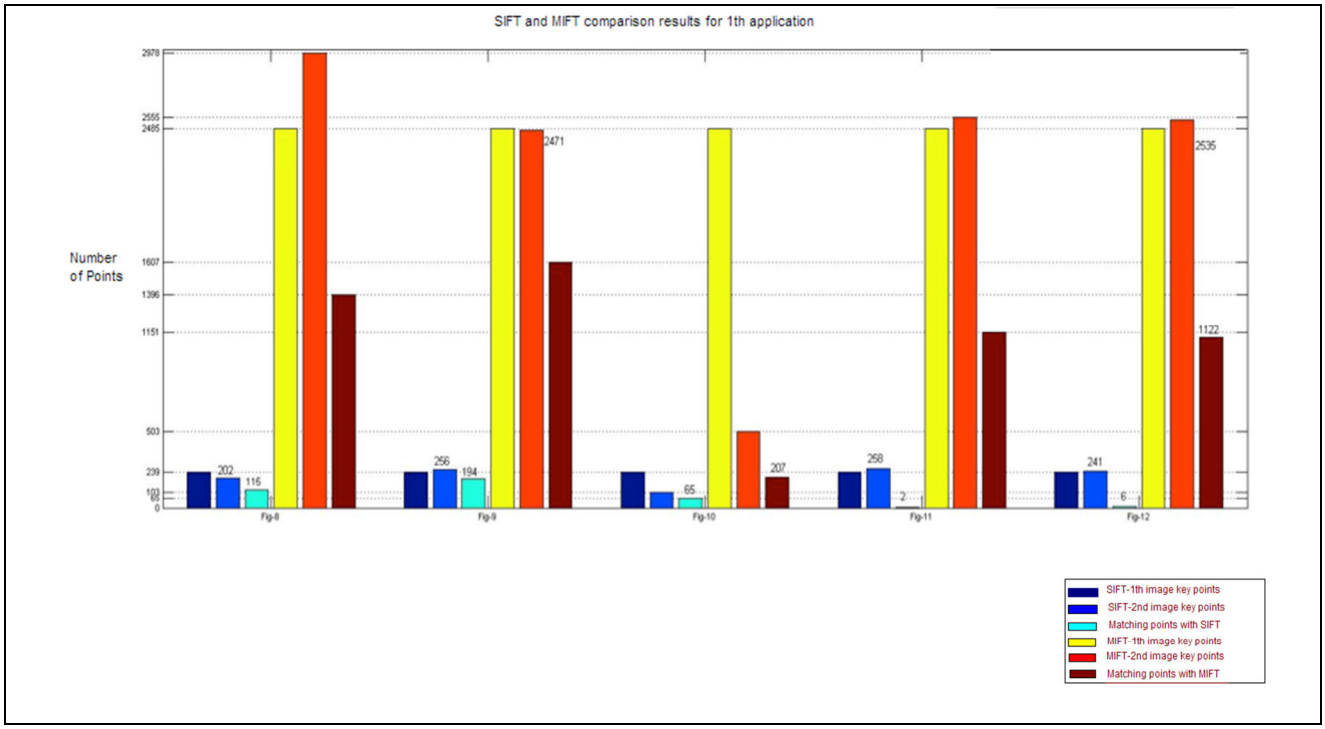


Figure-7: 5 different matching results from App-1 with using SIFT and MIFT methods and the key points and the match point numbers

When we examined the bar chart (Figure 7); firstly, in the first application, SIFT method was used on the different images of the glass object, there have been found 5 different matching images. Later, with using these images, defining key points and obtaining their numbers were the next step. In this diagram, number of key points for the 5 images was given as a navy blue bar in diagram and each of these was 239. For the 2nd application, all of the processes were made on the same images, there were found the number of key points again. In this diagram, this case was represented as dark blue in the bar chart and these values were 202, 256, 103, 258 ve 241, respectively.

For another case, all of the average points were calculated from the images with using SIFT method. In this diagram, number of key points for the 5 images was given as light blue bar in car chart and these values were shown as 115, 194, 65, 2 and 6, respectively.

Then, two applications were implemented with MIFT method and key points and rotations were obtained. In the diagram, bars that were yellow, presented number of key points for the first image and for all of the images, these values was 2495. Bars that were red, showed number of key points for the first image and for all of the images, these values were 2978, 2495, 503, 2555 ve 2535, respectively.

As the same,in detail, in Figure 8; For the first image with using SIFT method and for the second image with using SIFT method, number of key points, number of matching points were given in this bar graph. Also, for the first and second image with using MIFT method, number of key points, number of matching points were given in this bar graph again.

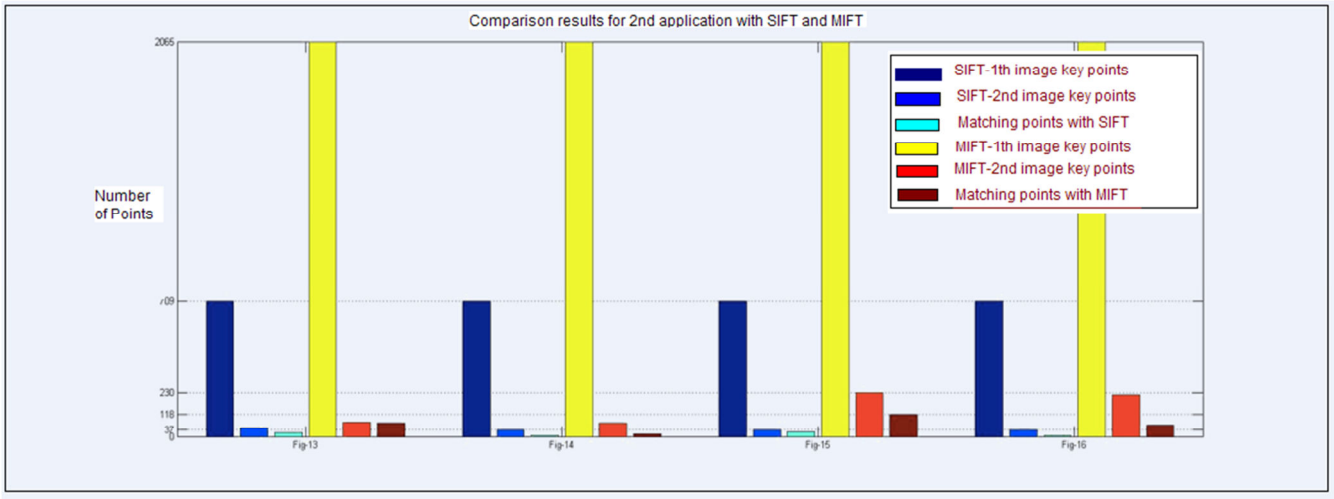


Figure-8: 4 different matching results from App-2 with using SIFT and MIFT methods and the key points and the match point numbers

If the bar chart which was given above was examined in detail; for the 2nd application, firstly, SIFT method was applied to the image and to its part and then 4 different matching results were obtained. Later, with using these images, all of the key points were calculated practically. In this diagram, for 4 different images, navy blue charts presented the total number of key points and this value was 709. Then, for the 2nd application, all of the processes were the same and in this case key points were identified. For 4 different images, dark blue charts were presented the total number of key points and this value was in the range of [0,37].

After that, two applications were implemented with the MIFT method again and for Application-2, all key points and rotations were found in results. In the diagram, yellow colored bars showed the 1st image's key points and for 4 different images, this value was 2065. Also, red colored bars showed the 2nd image's key points and for 4 different images and the values were, 77, 77, 230, 220, respectively.

When we examined Figure 7 and Figure 8, according to the results, there was an abrupt change in matching results. It was seen that, according to MIFT method, in mirroring processes, MIFT method had more power and resistance than SIFT method. To compare the scan time values of the methods, Figure 9 was presented practically and also process time for each of the key points were shown in this graphic.

Figure-9: Statistical representatiton of matching time for SIFT and MIFT method

This graphic was prepared with MATLAB Image Processing Toolbox. Horizontal axis showed the image data and the vertical axis showed the scanning time. According to the results, MIFT method was a time saver method and this has a lot of advantages. As seen, if filtering was performed powerfully, more key points were eliminated and there occured more stable new points.

REFERENCES

[1] Lowe, D.G., Object Recognition from Local Scale-Invariant Features. ***Proc. of the International Conference on Computer Vision***, (2): 1150-1157, 1999.

[2] Mikolajczyk, K. Schmid, C., "A performance evaluation of local descriptors". ***Dept. of Eng., Oxford Univ.***, 78-82 2005.

[3] Young, R. (1987). "The Gaussian derivative model for spatial vision: I. Retinal mechanisms". ***Spatial Vision2*** (4): 273–293(21)

[4] Davidson M. W., Abramowitz M., "Molecular Expressions Microscopy Primer: Digital Image Processing – Difference of Gaussians Edge Enhancement Algorithm", Olympus America Inc., and Florida State University.

[5] Krystian M., Cordelia S. "A performance evaluation of local descriptors", ***IEEE Transactions on Pattern Analysis and Machine Intelligence***, 10, 27, pp 1615-1630, 2005.

[6] Zhao W., Ngo C., Flip-Invariant SIFT for Copy and Object Detection.IEEE Transactions on Imageprocessing, 22: 3, 2013.

[7] Eski S., Recognition of Image Processing With Brand and Type Of Vehicle. ***Istanbul Technical University***, 2008.

[8] X. Guo, X. Cao, MIFT: A framework for featuredescriptors to be mirror reection invariant" ***Image and Vision Computing***, 30546-556 (2012).

[9] The Invariant Relations of 3D to 2D Projection of Point Sets, Journal of Pattern Recognition Research (JPRR), Vol. 3, No 1, 2008.

[10] "The Anatomy of the SIFT Method" in Image Processing On Line, a detailed study of every step of the algorithm with an open source implementation and a web demo to try different parameters

[11] Matthew J. McMahon, Orin S. Packer, and Dennis M. Dacey (April 14, 2004). "The Classical Receptive Field Surround of Primate Parasol Ganglion Cells Is Mediated Primarily by a Non-Gabaergic Pathway", ***Journal of Neuroscience***.

[12] Crowley, J, Riff O. "Fast computation of scale normalised Gaussian receptive fields", Proc. Scale-Space'03, Isle of Skye, Scotland, Springer Lecture Notes in Computer Science, volume 2695, 2003.

[13] Lowe, D. G. (2004). "Distinctive image features from scale-invariant keypoints". ***International Journal of Computer Vision*** 60 (2): 91–110.

[14] Lindeberg T. (1994). "Scale-space theory: A basic tool for analysing structures at different scales". ***Journal of Applied Statistics*** (Supplement on Advances in Applied Statistics: Statistics and Images: 21 (2). pp. 224–270

[15] Kim S., Yoon K., Kweon I., "Object Recognition Using a Generalized Robust Invariant Feature and Gestalt's Law of Proximity and Similarity", Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06), 2006

[16] ezSIFT: an easy-to-use standalone SIFT implementation in C/C++. A self-contained open-source SIFT implementation which does not require other libraries.

[17] Lowe, D. G., "Distinctive Image Features from Scale-Invariant Keypoints", ***International Journal of Computer Vision***, 60, 2, pp. 91-110, 2004.

[18] Liao K., Liu G.: An improvement to the SIFT descriptor for image representation and matching,10.1016/j.patrec.2013.03.021

[19] Soon T., Basari H., "Pattern Detection in Digital Image" 2010

[20] Ma R., Chen J., Su Z., "MI-SIFT: Mirror and inversion invariant generalization for SIFT descriptor," in Proc. Int. Conf. Image Video Retr., 2010, pp. 228–236.