Eye tracking by genetic deformable template matching algorithm in video sequences

Randa Atta [1], Rehab F. Abdel-Kader [2] and Sheren Zakaria [3]

ABSTRACT

Eye detection and tracking has broad applications in human-computer interaction, behavioural analysis, computer graphics, security, and biology systems. In this paper, we propose innovative eye detection and tracking method in video sequences. The proposed method incorporates deformable template matching and genetic algorithm to overcome problems associated with varying eye sizes and orientations. The method presented does not require any complex mathematical calculations or prior knowledge about eye geometry. Experimental result shows that the proposed scheme can easily be implemented. The method was tested on the VidTIMIT database. The detection accuracy was 92.2%.

KEYWORDS: Genetic Algorithm, Template Matching, Eye Detection, Eye Tracking

1. INTRODUCTION

Eye detection and tracking is a fundamental task in computer vision, with broad applications in human-computer interaction, behavioral analysis, and computer graphics, such as monitoring human vigilance [1], gaze-contingent smart graphics [2], and assisting people with disability. Various methods for eye detection and tracking are reported in the literature and can be classified into three categories: Appearance based methods [3], Feature based methods [4], and template based methods [5, 6, and 7].Appearance based method detect eyes based on their photometric appearance and require a large amount of data for training the classifiers (e. g. neural network or the support vector machine). Feature-based methods explore the characteristics of the human eye to identify a set of distinctive features around the eyes. Features such as the limbs, pupil (dark/bright pupil images), and cornea reflections are common features used for eye localization. However, feature-based methods have difficulties with low contrast images. In template-based methods a generic eye model, based on the eye shape, is initially designed. Subsequently, template matching is used to search the image for the eyes. Canonical template matching is computationally expensive, needs high contrast images, works only with frontal view faces and it only matches a template of fixed size in a predefined orientation. In any arbitrary image the matched region does not always have the same size as of template. Therefore, the template size must be initially adjusted according to the matched region by using the deformable templates[8].

In deformable template matching an eye model is allowed to translate, rotate and deform to fit the best representation of the eye shape in the image. Then, the eye position can be obtained through a recursive process. Although this method can detect the eyes accurately, it requires high quality images and the initial estimate of eye location should be properly specified near the actual eye position. Furthermore, exhaustive search is computationally expensive and it is not suitable for real-time implementation .Al-Mamun et al.[8]proposed a novel approach for eye detection which exploits the flexibility of deformable templates and uses a genetic algorithm to match the template for eye detection. The utilization of the genetic algorithm reduces the time required for template matching than the conventional template matching algorithm.

Most of the schemes described above for eye detection use prior heuristics about eye geometry and some of them really based on very complex mathematical calculations which are almost not possible for real-time implementation. While these methods have been successful in improving eye detection and tracking, reliably detecting and tracking eyes in conditions of variable face pose remains largely problematic. In this paper we propose a novel algorithm for eye detection and tracking in video sequences. The proposed method utilizes a threshold value for the correlation coefficient to switch between regular template matching and deformable template matching with GA. The proposed method was found to yield excellent recognition results while optimizing the execution time

The paper is organized as follows: in Section 2, template matching and genetic algorithms are explained. In Section 3, the proposed eye detection and tracking method is described. Section 4 reports the experimental results. Finally, we summarize the paper with some concluding remarks in Section 5.

[1] Assoc. Prof, Electrical Engineering Department, computer and control Section, Faculty of Engineering, Port Said University, Port Said, Egypt ,E-mail r.atta @eng.psu.edu
[2] Assoc. Prof, Electrical Engineering Department, computer and control Section, Faculty of Engineering, Port Said University, Port Said, Egypt , E-mail: rehabfarouk @eng.psu.edu
[3] Demonstrator, Electrical Engineering Department, computer and control Section, Faculty of Engineering, Port Said University, Port Said, Egypt, E-mail: sherenzakaria @eng.psu.edu

2. TEMPLATE MATCHING BY GENETIC ALGORITHM (GA)

2.1.Template matching

Template matching is a technique used in classifying objects by comparing portions of images against one another. It is a popular method for pattern recognition. Detailed description of this method can be found in [6, 7].

Definition 1: Let *I* be an image of dimension $m \times n$ and *T* be another image of dimension $p \times q$ such that $p<m$ and $q<n$ then template matching is defined as a search method which finds out the portion in *I* of size $p \times q$ where *T* has the maximum cross correlation coefficient with it.

2.1.1. Cross-correlation Coefficient

The normalized cross correlation coefficient is defined as:

$$\gamma(x, y) = \frac{\sum_s \sum_t \delta_{I(x+s,y+t)} \delta_{T(s,t)}}{\sum_s \sum_t \delta^2_{I(x+s,y+t)} \sum_s \sum_t \delta^2_{T(s,t)}} \quad (1)$$

Where: $\delta_{I(x+s,y+t)} = I(x+s, y+t) - \bar{I}(x, y)$
$\delta_{T(s,t)} = T(s, t) - \bar{T}$

$s \, \varepsilon\{1,2,3,.... ,p\}$, $t \, \varepsilon\{1,2,3,.... ,q\}$, $x \, \varepsilon\{1,2,3, ... ,m-p+1\}$, $y \, \varepsilon\{1,2,3, , ... ,n-q+1\}$

$$\bar{I}(x, y) = \frac{1}{pq} \sum_s \sum_t I(x + s, y + t)$$

$$\bar{T}(x, y) = \frac{1}{pq} \sum_s \sum_t T(s, t)$$

The value of cross-correlation coefficient γ ranges in [-1, +1]. A value of +1 indicates that there is a complete match between *T* and *I(x, y)* and a value of -1 indicates complete disagreement. In the Template matching algorithm the template *T* slides over image I and the value of γ is calculated for each position *(x, y)*. After completing this calculation, the point which exhibits maximum γ is referred to as the match point.

2.1.2. Deformable Template matching

The traditional template matching procedure explained above matches a template of a fixed size in a predefined orientation. But this is not always true, as in eye tracking in video sequences the eyes may be rotated or at a different scale other than in the template. Deformable template matching algorithm is utilized to address the difficulties associated with changes in eye scale and orientations.

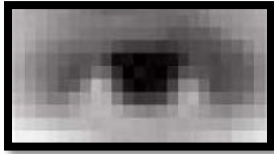The eye template used in this paper is presented in Figure1.

Figure 1.Template used in the experiment

In deformable template matching the eye template is deformed with various scales and rotation angles. The deformed templates are correlated with the face image. The region of the face which gives maximum correlation refers to eye region.

Exhaustive search in the deformable template matching algorithm is very time consuming. Therefore, to achieve real-time performance, the deformable template matching algorithm is combined with genetic algorithm.

2.2 Genetic Algorithm

Genetic Algorithm (GA) is a search technique used in optimal or near optimal solutions to optimization and search problems. GA uses techniques inspired by evolutionary biology such as inheritance, mutation, selection, and crossover. The algorithm begins by creating a random initial population then creates a sequence of new populations. To create a new population; Score each member of the current population by computing its fitness value. Scales the raw fitness scores to convert them into a more usable range of values. Selects members, called parents, based on their fitness. Some of the individuals in the current population that have lower fitness are chosen as elites. These elite individuals are passed to the next population. Children are produced from the parents either by making random changes to a single parent (mutation) or by combining the vector entries of a pair of parents (crossover).The current population is replaced by the generated children to form the next generation. The algorithm stops when one of the termination criteria is met. Detail study and application of genetic algorithm can be found in [8, 9 and10].

In our proposed method correlation of eye template with various overlapping regions of the face image is found out by changing eye template rotation angles and scaling factors using GA to avoid exhaustive blind search. The region with maximum correlation with the template corresponds to the detected eye region

2.2.1 Fitness function

The selected fitness function is the Cross correlation Coefficient calculated from (1).Therefore, the optimization problem to be solved with GA is to find the best rotation angle and scaling factor to maximize the fitness function. A fitness value closer to 1is a clear indication of a better template match. The region with the maximum correlation with the template refers to the detected eye region.

2.2.2. Chromosome Structure

Each chromosome is a vector of genes representing a trial solution.. Each gene can be a binary number, a real number or other symbols. In this Paper the used genes are coded as real numbers consisting of the scaling factor(s) and the rotation angle (φ) as shown in Figure2.The rotation angle is Confined in the range [-32,

+31]and the Scaling factor is confined in the range [0.5, 2.5].
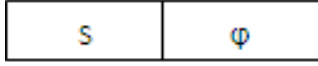

Figure.2. Chromosome structure

2.2.3 GA with population inheritance

In video processing, each frame is treated separately. The GA population is randomly initialized for every frame and this significantly increases processing time. To avoid this problem, the genetic information that has evolved in previous frames is reused in future frames. In fact, the initial population for a new frame is generated with the final population achieved in the previous frame. This population inheritance property can significantly reduce the processing time and increase the accuracy
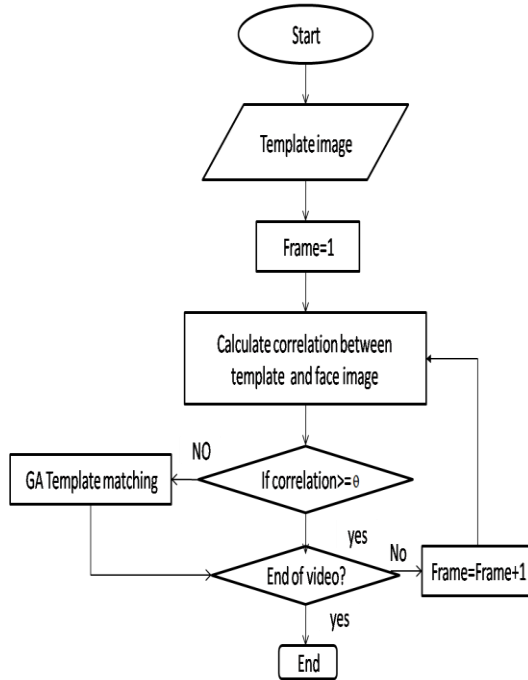
3. THE PROPOSED METHOD

Figure3.proposed method

Even with GA population inheritance we also need to reduce the processing time. The proposed method starts with the calculation of the cross-correlation coefficient between the eye template shown in Figure 1, and the face image. The maximum correlation value is calculated. If it is greater than or equal or a predefined threshold value (θ) then the point which exhibits maximum γ is referred to as the eye position. If the maximum correlation value is less than θ then the template needs to be deformed using the genetic based deformable template matching to find the eye position. This method can yield good detection accuracy while significantly reducing the processing time. The flowchart of the proposed method is presented in Figure 3.

4. EXPERIMENTAL RESULTS

The proposed eye tracking algorithm was implemented using MALAB 7.0 Software. The VidTIMIT database [11] was used to test the performance of the proposed method. VidTIMIT is comprised of videos and corresponding audio recordings of 43 people, reciting short sentences. Each person performed a head rotation sequence in each session. The sequence consists of the person moving their head to the left, right, back to the center, up, then down and finally return to center. The recording was done in an office environment using a broadcast quality digital video camera.

GA parameters were set as follows: population size is 10, probability of crossover is 0.7, probability of mutation is 0.05, and the type of crossover is uniform crossover. The population size is small, because this method will be applied for real-time processing in the future. The GA processing is terminated after 100 generations. To ensure result consistency five runs were performed for each video .

The estimated eye locations are compared with the exact eye center locations based on a relative error measure proposed by Jesorsky et al. [12], Let Cr and Cl be the exact eye center locations (by manual annotation), and Cr′ and Cl′ be the estimated eye positions. The relative error of this estimation is measured according to the formula:

$$\mathbf{d} = \frac{\mathbf{max}(\|\mathbf{Cr}-\mathbf{Cr}'\|,\|\mathbf{Cl}-\mathbf{Cl}'\|)}{\|\mathbf{Cr}-\mathbf{Cl}\|} \quad (2)$$

In a human face, the width of a single eye equals to the distance between inner eye corners. Therefore, half an eye width approximately equals to a relative error of 0.25[13]. Thus, in this paper, we considered a relative error of d < 0.25 to be a correct estimation of eye positions. Our method has a 92.2% overall success rate for d < 0.25 on the VidTIMIT database. Some examples of estimated eye locations are shown in Figure 4.

Figures 5, 6 present the calculated accuracy and the corresponding processing time for GA with and without population inheritance .It was found that GA with population inheritance has better accuracy and less time than GA without population inheritance. but we also still need to reduce time. Then our proposed method was tested for different values of θ .Figures 7, 8 present the calculated accuracy and the corresponding processing time for different values of θ.It was found that best value for θ to be .95with an average accuracy of92.2% and average processing time 15.2 sec per frame. These results showed better accuracy and less time than GA even with population inheritance . This indicate that the proposed method is effective.
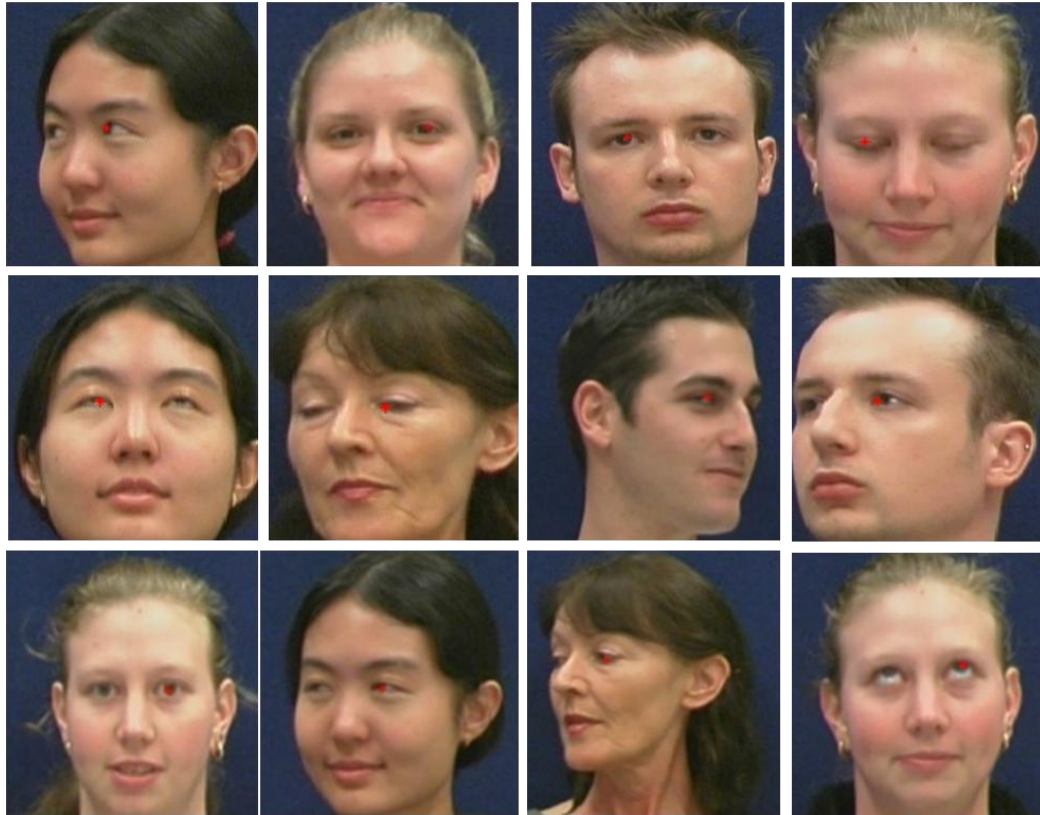
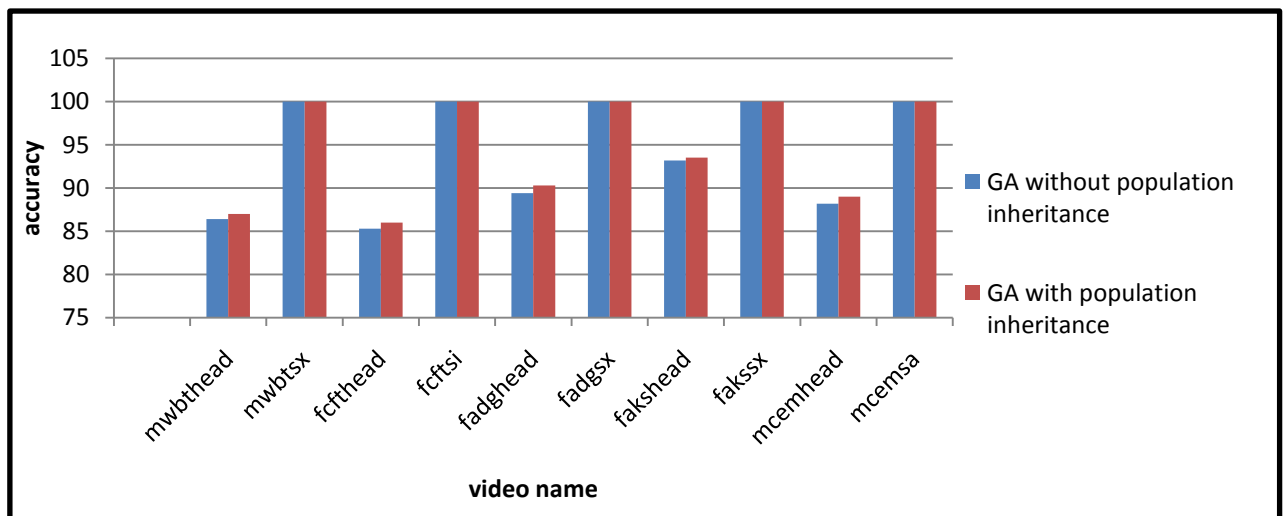133

Figure 4. Successful eye tracking

Figure.5. Accuracy comparison between GA and GA with population inheritance
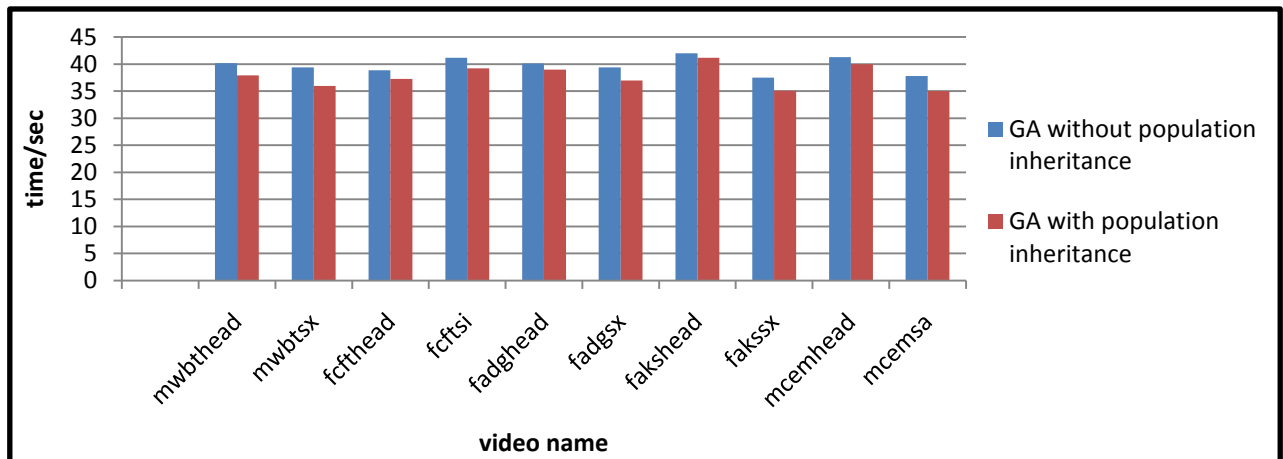
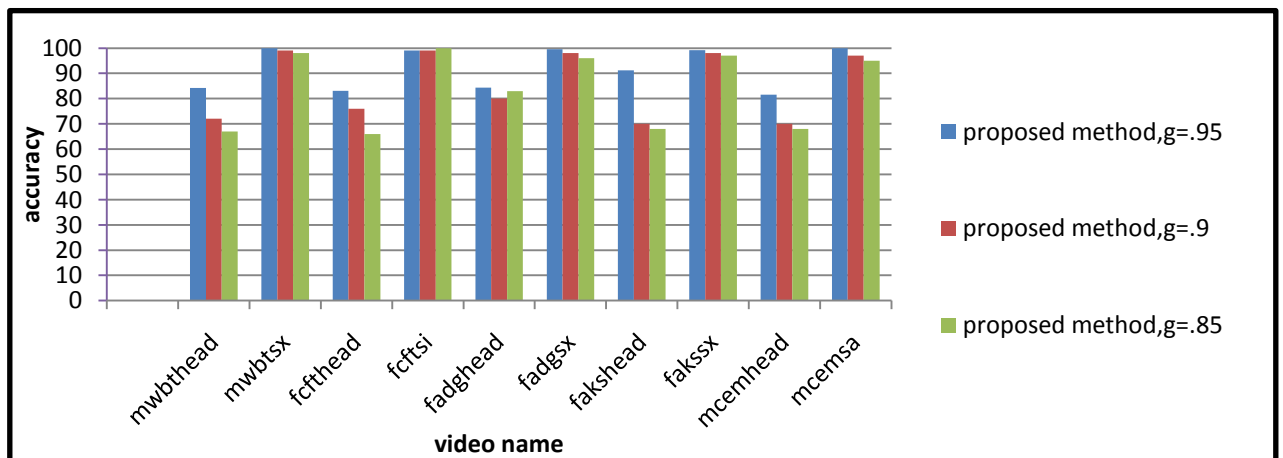**Figure.6. time comparison between GA and GA with population inheritance**

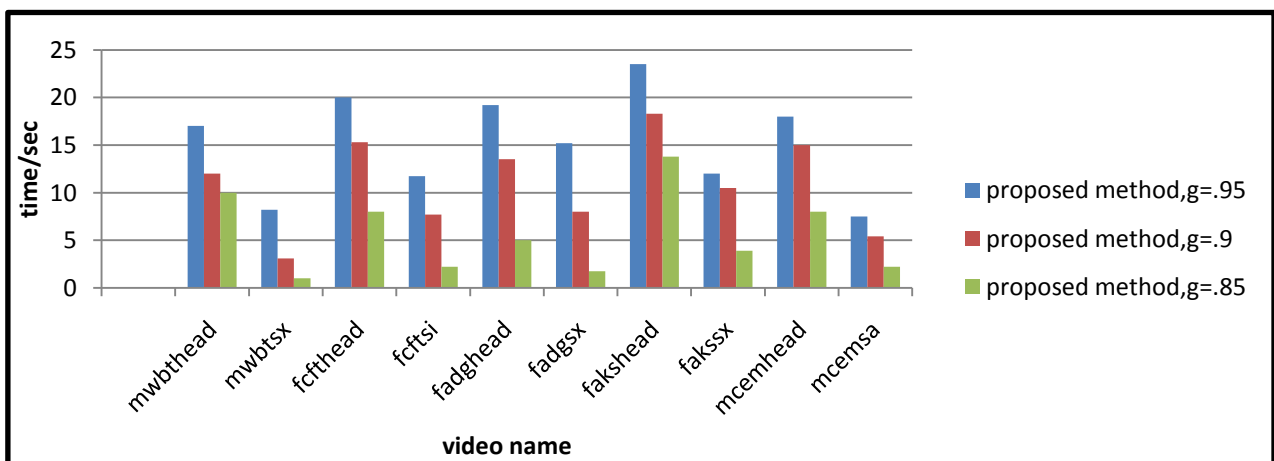**Figure.7. Accuracy comparison for different θ**

Figure.8. time comparison for different θ

5. CONCLUSION

In this paper new eye detection and tracking scheme is proposed. The proposed algorithm relies on genetic algorithm based deformable template matching. it incorporates deformable template matching and genetic algorithm to overcome problems associated with varying eye sizes and orientations The method does not require any complex mathematical calculation or prior knowledge about the eye geometry or potential eye location tags like face rotation angel, head contour or nose position. using a threshold value for the correlation coefficient to switch between regular template matching and deformable template matching with GA. which is a simple method and can be implemented easily by hardware. Proposed eye tracking in video system can find out eye location with success rate 92.2% and average processing time 15.2 sec per frame. In our future work, we will try to reduce average processing time and we also must compare to the other work with same dataset.

REFERENCES

[1] Qiang Ji and Xiaojie Yang," Real-Time Eye, Gaze, and Face Pose Tracking for Monitoring Driver Vigilance", Real-Time Imaging, Vol. 8, No. 5 London, UK, and UK: Academic Press Ltd, pp. (357-377), 2002.

[2] Qiang Ji and Zhiwei Zhu, "Eye and gaze tracking for interactive graphic display", 2nd International Symposium on Smart Graphics, Hawthorne, NY, USA, 2002.

[3] S.V. Sheela and P.A. Vijaya "An Appearance based Method for Eye Gaze Tracking ",Journal of Computer Science 7 (8),pp.(1194-1203), 2011

[4] Roberto Valenti, ZeynepYucel and Theo Gevers "Robustifying Eye Center Localization by Head Pose Cues " IEEE, pp .(612-618) ,ISBN 978-1-4244-3992-8, 2009

[5] Zhiwei Zhu, Qiang Ji, "Robust real-time eye detection and tracking under variable lighting conditions and various face orientations" Computer Vision and Image Understanding, Vol. 98, No. 1 pp.(124-154),2005.

[6] Nilamani Bhoi ,Mihir N. Mohanty."Template Matching based Eye Detection in Facial Image", International Journal of Computer Applications (0975 – 8887) Volume 12– No.5, December 2010

[7] R.M.Dufour, E.L.Miller, N.P.Galatsanos, "Template matching based object recognition with unknown geometric parameter" IEEE Transactions on Image Processing, vol. 11, issue. 12, pp. (1385 – 1396), 2002.

[8] Hawlader A. Al-Mamun, Nadim Jahangir, Md.Shahedul Islam and Md. Ashraful Islam," Eye Detection in Facial Image by Genetic Algorithm Driven Deformable Template Matching " IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.8, August 2009

[9] S.N.Sivanandam , S.N.Deepa," Introduction to Genetic Algorithms "ISBN 978-3-540-73189-4 Springer Berlin Heidelberg New York, 2008.

[10] Takuya Akashi, Yuji Wakasa, Kanya Tanaka, Stephen Karungaru, Minoru Fukumi, " Using Genetic Algorithm for Eye Detection and Tracking in Video Sequence", Journal of Systemics, Cybernetics and Informatics. Vol. 5 – No.2

[11] [itee.uq.edu.au/~conrad/vidtimit/] VidTIMIT database C. Sanderson and B.C. Lovell. " Multi-Region Probabilistic Histograms for Robust and Scalable Identity Inference " Lecture Notes in Computer Science (LNCS), Vol. 5558, pp. (199-208), 2009

[12] Oliver Jesorsky, K. J. Kirchberg, and Robert Frischholz. " Robust face detection using the hausdorff distance " . In Proc. International Conference Audio-and Video-Based Biometric Person Authentication (AVBPA), Springer, Lecture Notes in Computer Science, LNCS-2091, pp. (90–95), 2001.

[13] Mehmet Türkan, Montse Pardás, A. Enis çetin, "Human eye localization using edge projections ". VISAPP 2007: Proceedings of the Second International Conference on Computer Vision Theory and Applications, Barcelona, Spain, Vol.1,2007