Proceedings of the 10th
World Congress on Intelligent Control and Automation
July 6-8, 2012, Beijing, China

Fast Eye Localization Based on a New Haar-like Feature*

Yefei Chen, Jianbo Su

Department of Automation
Shanghai Jiao Tong University
Shanghai, China

{marschen, jbsu}@sjtu.edu.cn

Abstract **- This paper focuses on fast eye localization method. According to the priori proportional relationships of face features, we firstly set an appropriate candidate window from the face region detected. Secondly, histogram equalization is applied on the candidate region to eliminate illumination effects. This paper presents a new haar-like feature generating the confidence of the feature throughout the candidate region in order to locate eyeball accurately and rapidly. Our method is proved to be simple and robust against the disturbance caused by glasses, eyebrow and hair. The process of training and learning is not necessary because of the appropriate priori knowledge. Our experiment on three face databases shows that our method can be applied to real time eye localization and even to pupil localization under most circumstances, achieving accurate results.**

Index Terms – Face detection, Face recognition, Eye localization, threshold, haar-like feature.

I. INTRODUCTION

Generally, face recognition is composed of three parts: face detection, feature extraction and feature comparing [16]. Image rotation and normalization, which are based on the operation of eye localization, are vital to the feature extraction and follow-up steps [1]. Because of this, the research of eye localization is still in the ascendant, which involves many variant state-of-the-art methods including region segmentation [2], edge extraction [3], grayscale projection [4], template matching [5] and so on.

The effect of region segmentation is very rough caused by the disturbance of eyeglass. The fact of edge extraction is to find an eye template using Hough transfer, which demands large pre-process. Still, even some improved methods based on that will confront the disturbance caused by eyeglass and eyebrow. Grayscale projection is a kind of rapid algorithm, locating eye position according to the peaks and valleys of the projection on two different axes of coordinate. Though projection peak focuses on the diversification of the grayscale levels, black-frame glasses, eyebrow and hair are still hard to be distinguished depending on the two-dimensional projection. Template matching needs to normalize the scale and orientation of face image, which costs expensive computation [14], while the templates are also obtained through training. Among these algorithms, Adaboost algorithm [7] based on sample training has a significant effect in the aspect of eye localization which strictly demands abundant training samples. However, eyebrow with high grayscale levels may be defined as eyeball in practice. Especially, the scale of candidate searching window is constraint to the size of training samples which may cause low recognition rate on low resolution frames.

In terms of the problem of eye localization, taking advantage of the distinct characteristic of the eyeball is key to eye localization no matter what method is used. Intuitively, eyeball has a shape of ball though in most condition it is blocked which makes it a non-round shape. In the aspect of grayscale, the values of grayscale near the pupil are obviously different. In the aspect of position, the eyeball lays on the upper part of the face. Still, obstacles which caused by eyeglass, eyebrow, hair and variable illumination are inevitable. It is clearly stated that how to take full advantage of the distinct characteristic of eyeball is of great importance to solve the problem.

According to the issues mentioned above, this paper proposed a new localization algorithm based on histogram threshold which shared the principle of integral projection algorithm [6][8]. First, binary transform is performed on the candidate region by a proper threshold and then the fast searching algorithm is applied on the same region. With regard to the geometry of the eyeball, this paper analyzes the relationship of shape and location between disturbance and object and presents a new kind of haar-like feature based on basic haar-like features [12]. Classical haar-like features are presented by P. Viola et al. in [7] 2001, which is applied to boosted cascade to detect object rapidly. Later, R. Lienhart et al. improved the feature and extended to 3 kinds of haar-like features: edge feature, line feature and center-surround feature in [12]. These extensions enable the detection to be more robust against the changes of pose. The main principle is extracting the features from the pixel-level via the difference of the sum of pixels in the rectangles nearby. Intuitively, rectangle features can reflect the specific characteristic of the object such as that region near the eyeball is darker than that near the nose and cheek. However, eyeball possesses distinct features different from other objects. What's more, our algorithm is applied after threshold segmentation which is different from the regular haar-like features which extract the features from the raw pixels. A new haar-like feature which is used to search the candidate region and to get a matrix of confidence is proposed in this paper. Finally, our criterion is presented to analyse the confidence matrix and locates the

* This work is partially supported by National Science Foundation of China Grant #60935001.

position of eyeball. The later experiment shows that our algorithm can locate the eyeball rapidly and accurately.

II. Eye localization based on histogram

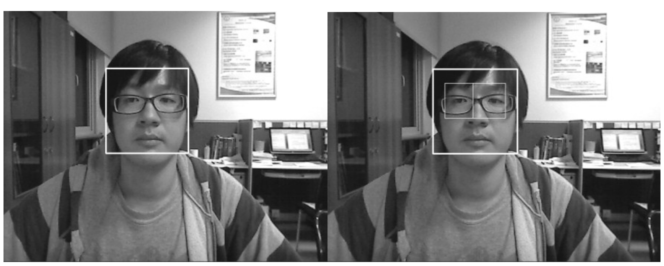
This paper researches the method to locate the eyeball rapidly upon the face images detected previously. Approaches to obtain the face region of interest are various [9], such as Adaboost algorithm which is used to obtain the face region for eye localization in this paper (Fig. 1(a)).

In order to locate the eyeball, we firstly take the advantage of geometric relationships between face and eyeball such as that the eyeball locates on the upper part of the face regularly. With this priori knowledge, the candidate region for location can be narrowed (Fig. 1(b)). Secondly, threshold segmentation is used to segment the region containing the eye because of the fact that the grayscale levels near the pupil are deeper than those far away from the pupil [15]. Apparently, disturbances caused by hair, eyeglass or eyebrow are inevitable judging from Fig. 1(b).

A. *Threshold and Segmentation*

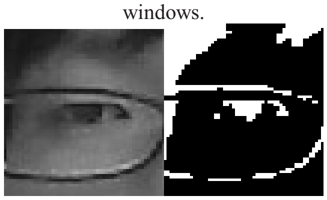
Threshold segmentation is a widely applied image processing method, of which the key point is the selection of threshold. In this paper, histogram analysis is applied to select a proper threshold. According to the histogram of the candidate region, the last p% pixels with low grayscale levels are set to 255 and others 0. Though this method can select the threshold rapidly and segment the region containing the eyeball effectively, it theoretically depends on the value of p to some extent.

In order to eliminate the variable illumination, histogram equalization is applied to the candidate region before being binarized (white region in Fig. 2(b)). Other method of image processing (LTV) is also used to improve the result later.



(a) (b)

Fig. 1(a) Effect of face detected. 1(b) Effect of face detected with candidate windows.



(a) (b)

Fig. 2(a) Candidate region of eye. 2(b) Result of threshold segmentation.

B. *New kind haar-like feature*

From Fig. 2(b), the eyeball region shapes like a round, where the white pixels are density ideally after binary, which is similar to the centre-surround haar-like feature in fact. However, the haar-like feature mentioned in [12] is over complete, which needs the process of training to select the proper representative features. Because of not being normalized in scale, that feature can hardly be applied to locate eyeball directly.

In terms of frontal face, the disturbances caused by hair, eyebrow and eyeglass have strong continuity in vertical or horizontal direction separately. According to the concept from the gradient, eyebrow has strong continuity in x-direction, while frame of eyeglass has strong continuity in x-direction or y-direction with different area and hair appears from the border of the candidate region into the region. Hence, we present the new kind haar-like feature which generates the feature value D by calculating the difference of the sum of the pixels in four white surrounding rectangles between the sums of the pixels in the centre rectangle illustrated in Fig. 3. When the eye region falls on the regional center, the feature will reach the maximum within the candidate region. Hair will increase the value of D2 and D4, while eyebrow will increase the value of D1 and D3. The difference from the centre-surround features is that our feature does not take the region in 45 degrees direction into consideration because the main disturbance comes from x or y direction.

According to (1) and (2), the confidence C of feature will be generated, where W stands for the weights of different rectangle of the feature, I(x, y) stands for 1 when the pixel (x, y) value is 255 and 0 otherwise and S(x, y) stands for 1 for each pixel. With the consideration of four rectangle density with different directions, the confidence C presented below can be used to locate the eyeball against the interference caused by eyeglass, eyebrow and hair.

$$C = W_0 \cdot D_0 + W_1 \cdot D_1 + W_2 \cdot D_2 + W_3 \cdot D_3 + W_4 \cdot D_4 . \quad (1)$$

$$D_i = \sum_{x,y} I(x, y) \Big/ \sum_{x,y} S(x, y) , \quad (i = 1, 2, \ldots, 4). \quad (2)$$

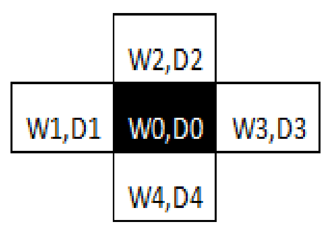


Fig. 3 New haar-like feature

C. *Algorithm based on new kind haar-like feature*

When searching all the windows throughout the candidate region with the new kind of haar-like feature， a matrix of feature confidence $C_{i,j}$ (where i, j stands for the position of the centre of the feature) will be formed, of which each value can be computed by (1). Then, coordinate the position of the eyeball based on the confidence matrix. The confidence matrix stands for the confidence map of the eye candidate region as illustrated in Fig. 4. Apparently, the position of maximum in the confidence map is the eye location, which can be computed through (3). But simple solution of maximum is easy to get local minima. Therefore, we introduce the follow-up criteria avoiding such occasion.

$$C_{I,J} = \arg\max_{i,j}(C_{i,j}). \tag{3}$$

In order to overcome the disturbance caused by eyebrow and hair, we consider that 1. The confidence of eyeball will be no less than 80% of that of eyebrow. 2. Eyeball must locate under the eyebrow. 3. Hair must extend to the candidate window from the boundary. Hence, on the basis of the views mentioned above, we establish the follow-up method:

1) Expand the boundaries of eye candidate box. Set the grayscale of pixels in the expanded region to 255 which will decrease the confidence value covered by hair and eliminate the interference of hair.
2) Find the maximum $C_{i,j}$ in matrix $C_{I,J}$ using equation (3). Then set the surrounding region, which is 1/25 large of the candidate region, around the maximum to be the minimum in matrix $C_{i,j}$ in order to avoid falling into extreme value.
3) Using (3) to find the new maximum $C_{P,Q}$.
4) If $C_{P,Q} \geq 0.8 \times C_{I,J}$ and J < Q, then (P, Q) is the position of eyeball. Otherwise, (I, J) is the location of eyeball.

Because the haar-like feature itself has the characteristic of fast calculation, applying the feature and criteria presented in this paper can rapidly and accurately locate the eyeball. The basic algorithm flow chart is illustrated in Fig. 5.

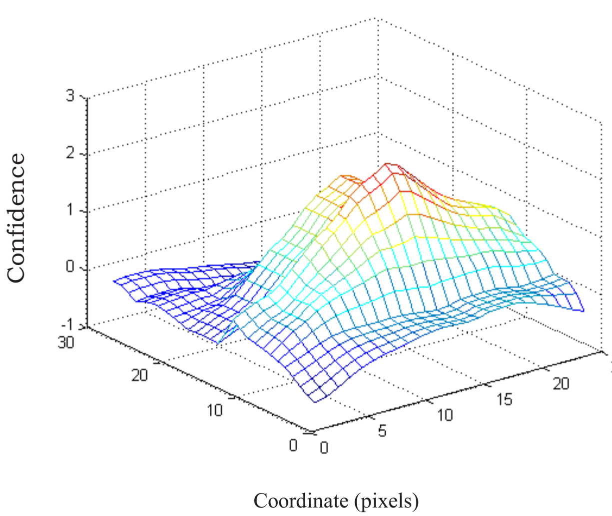


Fig. 4 Confidence map of matrix C

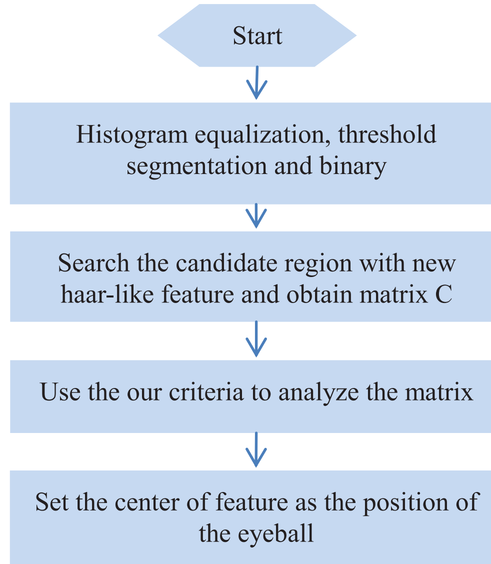


Fig. 5 Flow chart of the eye location algorithm

III. EXPERIMENTS AND ANALYSIS

In order to exam the stability of the algorithm, a widely used accuracy evaluation of eye location proposed in [11] is accepted in this paper. Suppose that C_l and C_r are eye positions calibrated by hand, $C_l^{'}$ and $C_r^{'}$ are eye positions located by algorithm, d_l is the Euclid distance between C_l and $C_l^{'}$ and d_r is the Euclid distance between C_r and $C_r^{'}$, the relative error of eye position is :

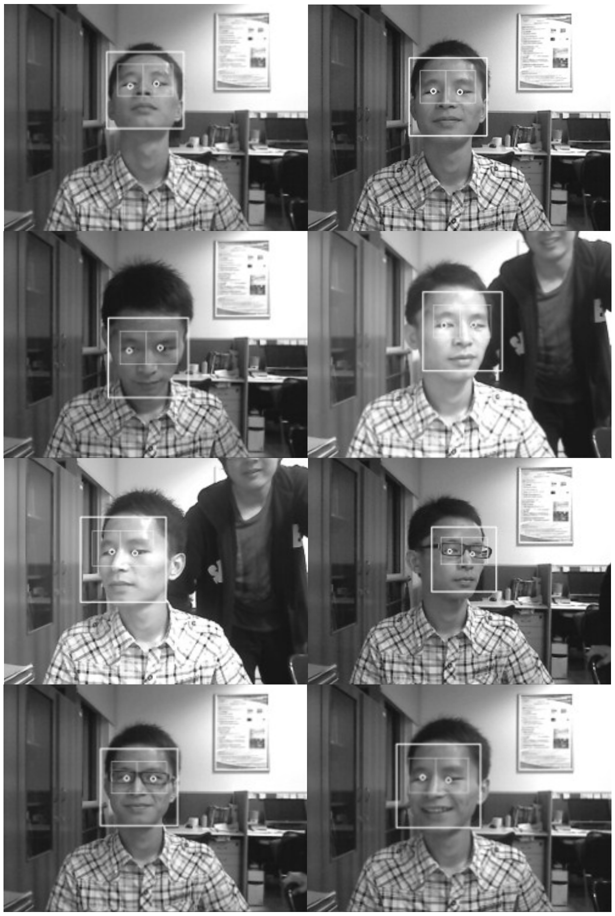
$$err = \max(d_l, d_r) / \left\| C_l - C_r \right\|. \tag{4}$$

If err<0.25, then the localization is considered correct. Therefore, the localization accuracy based on the face data containing N images is (where err_i stands for the relative error of the ith face image.):

$$rate = \sum_{\substack{i=1 \\ err_i < 0.25}}^{N} 1/N \times 100\%. \tag{5}$$

A. Real time experiment

The method, based on haar-like feature, presented in this paper has the characteristic of rapid computation. Our algorithm is real–time examed on PC with Intel P4 1.8G CPU and 1G memory using camera with resolution 640*480. The Adaboost algorithm is used to detect face and segment face region of interest. The real-time eye windows are set as follow-up data: assuming that width and height of the face region is W and H, the origin of left eye window is (W/7, H/6) and the origin of the right eye window is (W/2, H/6) with both (W/3, 5*H/12) large. The real time result is showed in Fig. 6 (p=10), which contains images with different poses, distances. It is proved that our algorithm is effective to tilted face to the extent of $\pm 30^{\circ}$ by calculating the eye coordinates. Also, faces within ±20° are located as Fig. 6 demonstrated. Especially, images illustrated in Fig. 6(a) contain some condition of eye closed and frames illustrated in Fig. 6(b) contain the interference caused by black frame eyeglass and hair. What is more, our experiment shows that area of face ranges from 300*300 pixels to 70*70 pixels all can be used to locate. Images with low resolution can be located properly by our algorithm which is important to real time localization.



(a)

(b)

Fig. 6(a) Results under variable poses. 6(b) Results under variable poses and distance including eyeglass.

B. *Experiments on standard face databases*

Standard face databases including FERET, JAFFE and BioID are used as probe set in this paper, containing 3816 calibrated images from FERET, 213 images from JAFFE and 1521 images from BioID. Face regions are segmented as the probe set for the experiment in advance.

If the width and height of face image segmented is W and H, then the origin of the left eye candidate window is (W/12, H/12), while the origin of the right eye candidate window is (W/2, H/12) with size of (5*W/12 , 5*H/12) each. The weight of haar-like feature is set as the follow-up data: W_0=4, W_1= -1, W_2= -1, W_3= -1, W_4= -1. Meanwhile, a constant scale of feature is used considering the cost of localization, which is set as 1/25 large of the candidate window. Images below are part of the samples used as probe set as illustrated in Fig. 7.

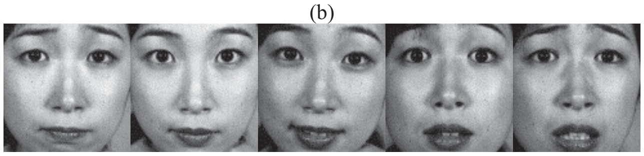
Firstly, the value of threshold p is assigned through these three face databases, which is showed in Table I in which the bold underlined part is the final result of the threshold. As we can see, when the threshold ranges from 5 to 15, our method can still locate the eyeball effectively which means that our method is not sensitive to the selection of p compared with [10]. That is to say, algorithm presented in this paper can be applied to other adaptive threshold pre-processing means. In our experiment, p is assigned to be 6 which shows satisfied accuracy rate according to the results shown in Table I.

The results of localization through our algorithm are shown in Fig. 9, which contain varying pose changes, illumination and reflection caused by eyeglass. It is proved that our method is robust to pose and accessory. Especially, in some closed eye situation, our method can still locate the eyeball accuracy. Meanwhile, take JAFFE for example. It is significant that our method can locate the eyeball accurately and even locate the position of pupil.

Besides histogram equalization, our method can be applied with different methods of image processing too. In order to alleviate the effect of illumination, we also apply LTV (logarithmic total variation model) [17] before eye localization to normalize the images from FERET illustrated in Fig. 8. The result illustrated in Table II shows that our method hardly depends on the value of p, which ranges from 2 to 10, after image pre-process. Also, the accuracy rate is a bit better than those using the raw images from the database.

(a)

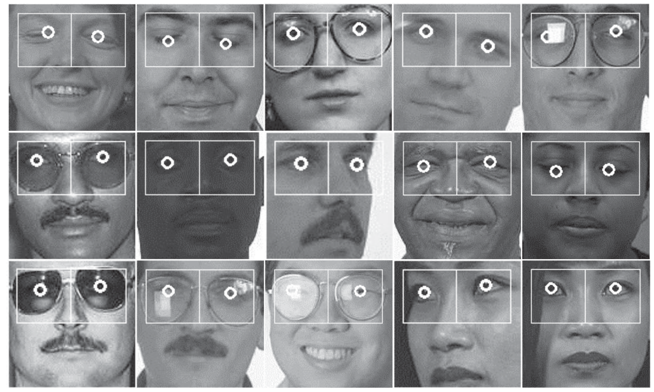
(b)



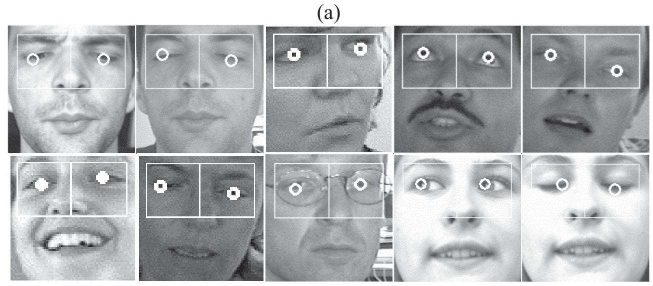
(c)

Fig. 7(a) Samples in FERET. 7(b) Samples in BioID. 7(c) Samples in JAFFE.

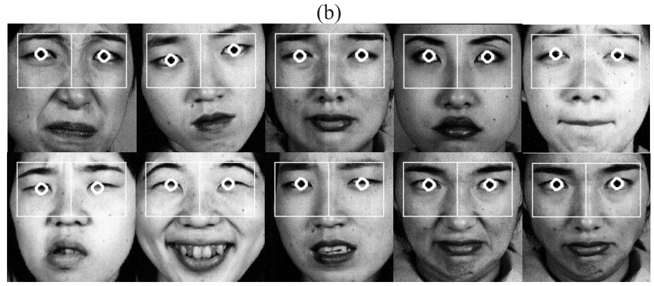
Fig. 8 Images after LTV transformation.



(a)



(b)



(c)

Fig. 9(a) Results on FERET. 8(b) Results on BioID. 8(c) Results on JAFFE.

TABLE I

ACCURACY RATE WITH VARIABLE P

p(%)	Accuracy rate		
	FERET (%)	BioID (%)	JAFFE (%)
3	98.68	93.95	62.44
4	98.66	95.00	83.56
5	98.71	95.72	97.65
6	**98.58**	**96.25**	**99.53**
7	98.29	96.71	100.00
10	96.85	95.72	100.00
15	93.55	95.13	100.00
20	87.05	91.05	100.00

TABLE II

ACCURACY RATE WITH VARIABLE P

P(%)	Accuracy on FERET(%)
2	98.71
3	98.71
4	98.79
5	98.74
6	**98.66**
7	98.50
10	97.98

C. *Comparison with other methods*

Other state-of–the-art algorithms are compared with our algorithm in this paper, which are listed in table III, among which the Adaboost algorithm presented by [7] is a widely used object detection method currently. We choose 4532 images with size 20*20 as positive eye samples with 2236 images each stage and finally obtain 16 cascades. The method in [8] is classical projection peak; the method in [10] is an improved projection peak; the method in [13] is an enhancement of VPF; the method in [14] is based on sampling training which using Adaboost and SVM together with HOG descriptor. According to Table III, our method is more accuracy than the methods without training and a bit less than those based on training.

Though our algorithm is less faster than projection peak, it has great improvement than methods, based on training, which is stated in [7][14]. Meanwhile, our algorithm is more accurate than projection peak. Our method is fast and simple. Even in long distance with frame 64*64, eyes can be located by our method while Adaboost algorithm demands frames with size upper 80*80. Theoretically, Adaboost algorithm focus on eliminating negative samples which cause that only half of the frames detected can be located comparing 90 per cent of our method.

TABLE III
ACCURACY COMPARED WITH DIFFERENT METHOD

Method	FERET (%)	BioID (%)	JAFFE (%)	Time(ms)
Our method	98.58	96.25	99.53	7
Method in [7]	98.93	96.03	99.47	10.3
Method in [8]	——	94.81	97.18	0.5
Method in [10]	98.92	95.87	99.26	0.56
Method in [13]	——	——	99.53	——
Method in [14]	99.6	——	——	161.8

IV. CONCLUSION

The haar-like feature proposed in this paper takes full advantage of human eye in the geometric and grayscale features on the characteristics different from eyebrow and eyeglass. Through the criteria upon the confidence matrix obtained by the feature, we find an effective solution to the interference of hair and eyebrow. However, under some circumstance such as eyeglass reflection and closed eyes with thick black eyebrow will yet give rise to the mistaken of localization. As demonstrated in our experiment, our method can locate eyeball effectively with variable pose changing ($\pm 20^{\circ}$) and rotation ($\pm 20^{\circ}$). Further research will focus on finding a more adequate threshold segmentation method in order to make the algorithm more robust.

REFERENCES

[1] S.G. Shan, Y.Z. Chang, W. Gao. Curse Of Mis-Alignment In Face Recognition: Problem And A Novel Mis-Alignment Learning Solution. Proceeding of the 6th IEEE International Conference on Automatic Face and Gesture Recognition, Korea, May, pp. 2004.

[2] S.G. Shan, W. Gao, X.L. Chen. Facial Feature Extraction Based on Facial Texture Distribution and Deformable. Journal of Software, 2001, vol. 12(4), pp. 570-577.

[3] X. Zhang, Y.J. Xu, L.M. Du. Automatic Localizing Main Facial Features in Color Images. Journal of Image and Graphics, 2000, vol. 5(2), pp. 128-143.

[4] J.Q. Feng, W.B. Liu, S.L. Yu. Eyes Location Based on Gray-level Integration Projection. Computer Simulation, 2005, vol. 22(4), pp. 75-76.

[5] H.R. Shi, X.S. Zhang, Y. Liang, H.C. Zhang, Y.M. Cheng. An Improved Template Matching Eye Location Method. Computer Engineering and Applications, 2004, vol. 33, pp. 44-45.

[6] G.H. Li, X.P. Cai and X.S. Li. An Efficient Face Normalization Algorithm Based on Eyes Detection. Proceedings of the 2006 International Conference on Intelligent Robots and Systems, Beijing China, October, 2006, pp. 3843-3848.

[7] P. Viola, M. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, Hawaii USA, December, 2001, pp. 511-518.

[8] Z.H. Zhou, X. Geng. Projection functions for eye detection. Pattern Recognition, 2004, vol. 37(5), pp. 1049-1056.

[9] Y. Ryu, S. Oh. Automatic Extraction of Eye and Mouth Fields from a Face Image using Eigenfeatures and Multilayer Perceptrons, Pattern Recognition. 2001, vol. 34, pp. 2459-2466.

[10] J.W. Dai, D. Liu, F. Yang, J.B. Su. Rapid Eye Localization Based on Projection Peak. Pattern Recognition and Artificial Intelligence, 2009, vol. 22(4), pp. 605-609.

[11] O. Jesorsky, K. Kirchberg, R. Frischholz. Robust face detection using the Hausdorff distance. In: J. Bigun, F. Smeraldi Eds. Lecture Notes in Computer Science 2091, Berlin: Springer, 2001, pp. 90-95.

[12] R. Lienhart, J. Maydt. An Extended Set of Haar-like Features for Rapid Object Detection. Proceedings of the International Conference on Image Processing, 2002, vol.1, pp. 900-903.

[13] M. Montazeri, H. Nezamabadi-pour. Automatic Extraction of Eye Field from a Gray Intensity Image using Intensity Filtering and Hybrid Projection Function, Communications, Computing and Control Applications (CCCA), 2011 International Conference on , March 2011, pp.1-5, 3-5.

[14] D. Monzo, A. Albiol, J. Sastre, A. Albiol. Precise Eye Localization using HOG Descriptors. Machine Vision and Applications, 2011, vol. 22(3): 471-480.

[15] K. P. Sushil, B. Saida, S. U. Mohammad. Automatic Adaptive Center of Pupil Detection Using Face Detection and CDF Analysis. Communications in Computer and Information Science, 2011, vol. 166, pp. 327-338.

[16] W. Zhao, R. Chellappa, R. J. Phillips, A. Rosenfeld. Face Recognition: A Literature Survey. ACM Computing Surveys, 2003, vol. 35(4), pp. 399-458.

[17] X.H. Xie, W.S. Zheng, J.H. Lai, P.C. Yuen. Face illumination normalization on large and small scale features, Computer Vision and Pattern Recognition(CVPR), 2008, pp.1-8, 23-28.