

A Real Time Robust Eye Center Localization using Geometric Eye Model and Edge Gradients in Unconstrained Visual Environment

Krupa Jariwala
Assistant Professor
Computer Engineering
Department
SVNIT, Surat

Aakash Nandi
Computer Engineering
Department
SVNIT, Surat

Upena Dalal, PhD
Associate Professor
Electronics Engineering
Department
SVNIT, Surat

ABSTRACT

Accurate eye localization is an essential and fundamental step in initialization for other eye tracking applications. A fast and accurate eye center localization method is proposed in this paper. A novel geometric eye model is derived based on face anthropometry parameters to identify coarse eye region correctly. A novel voting method using edge gradients on the iris boundary are used for fast and accurate eye center localization. A precise range around the boundary of an iris is derived to indicate the region within which the gradients are allowed to vote. Additionally the range between the pupil centers is derived to validate the search region. A weight map is generated for efficient computation, which is combined with edge gradients and the maximum of the multiplication between the dot products and the weight from the map is identified as the eye center. The proposed method is evaluated on challenging BioID database and found to be highly accurate for eye center localization task. The proposed method is efficient under natural lighting condition, has low computational complexity and excellent real-time ability.

Keywords

Eye center localization, Gaze Estimation, Human Computer Interaction.

1. INTRODUCTION

Eye localization is fundamental to all eye tracking applications such as eye gaze estimation in Human Computer Interaction(HCI), human cognitive analysis, website and product usability testing, educational training, driver assistance system etc. The task of eye localization is difficult due to variations in appearance, facial expression, occlusion, pose, illumination, imaging quality and motion characteristics of an eye and may lead to a complex system design.

There are many intrusive techniques available for eye center localization that makes use of electrooculography, Scleral coils, head mounted device with active infrared illumination. These techniques permit very accurate eye center localization however they have serious health concerns and are less efficient in daylight condition and outdoor situations. A vision based method is more appropriate in these situations however it is less accurate due to variations and appearance of an eye characteristic in real time scenarios. Most video based eye localization methods adopt two-stage analysis as face detection and eye localization, where the face is located first and the eyes are then detected by searching them within the possible eye region in the face. Correct face detection would reduce the search space and increase the accuracy of

eye detection. There are various face and eye detection methods broadly classified as following:

1.1 Face detection methods

A. Template matching methods : Here, the templates of various facial feature are constructed which are then compared with an input image. The eigenface [1, 2] is widely used template matching method where the eigenface of the input face image is compared with all the templates in the database to estimate the corresponding similarities and most matching one is decided by the similarity values. However, these methods are inefficient in case of affine transformation such as scaling, rotation, skew etc.

B. Appearance-based methods : These methods consider features of the whole face such as skin colour, face contour, facial feature etc. as a single feature does not provide enough information for the face detection. They also use various machine learning techniques to learn these features to classify the face patterns. Viola Jones[3] method based on integral image and Ada Boost cascade classifier performs well for face detection in most of the cases.

1.2 Eye detection methods

A. Shape based methods : The eyes can be described by their shape, which is composed of an iris, pupil, and the eyelids. The ellipse shape of an eye or circular shape of an iris is a frequently used feature in shape based methods. However these methods require high contrast images and may not work for dark iris. Another variation is to construct a deformable model [4] and use a template to identify the corresponding feature points of the eye, but these approaches are computationally complex.

B. Appearance-based methods : These methods detect eyes based on photometric appearance, including colour information and the surroundings of an eye such as template matching[5] or pattern matching based on Haar Like features[6,7]. These methods need large learning samples to provide accuracy to various deformations. These methods deal with rotations and scaling but fails if the hair or the glasses occlude the eyes.

2. RELATED WORK

Jesorsky[8] were among the first to highlight the performance and accuracy of the face detection. He proposed a measure for quantitative evaluation for the performance of Eye Localization methods thru Hausdorff distance (HD), which is widely used nowadays. The Hausdorff distance (HD) is a metric between two point sets that measure a relative error based on the distances between the expected

and the estimated eye positions to validate the performance of the eye detection system.

Recently, there have been many leading methods available for eye localization. They are based on Pairwise Reinforcement of Feature Responses (PRFR) and an Active Appearance Model (AAM) [11], Registration and Appearance Verification[12], 2D cascade classifier with Large training dataset[13], edge projection using wavelets[14], Isophote Curvature with MIC+SIFT+kNN Learning [15], cumulative density function (CDF) framework[16], Multiscale Sparse Dictionaries (MSD)[17], image gradients[18], E-LBP with probabilistic cascade (P-Cascade) framework[19], Scale-invariant feature transform (SIFT) and 2-class sparse representation classifier (SRC)[20]. The in depth review and performance analysis of state-of-the-art methods, based on their existing published results that adopt similar evaluation criteria for objective performance evaluation is available at [9][10].

3. PROBLEM STATEMENT

In uncontrolled scenarios, the geometrical structures and the appearance of an eye is drastically different from the predefined template or training data set and the accuracy of most of the eye localization methods decrease significantly. Therefore, in spite of an extensive research, the reliable eye localization in uncontrolled environment still needs to be addressed thoroughly.

The performance analysis of various methods based on Normalized Error Measurement for various error value of $e \in \{0.05, 0.10, 0.15, 0.20, 0.25\}$ is carried out at [10]. This measure indicate that for $e \leq 0.05$ the eye center is within diameter of the pupil, for $e \leq 0.10$, it is within diameter of the iris and for $e \leq 0.25$, it is within the eye center and the eye corners. For eye center localization task, a good performance of $e \leq 0.05$ is highly essential. However, higher values ($e \leq 0.25$) are acceptable for other general applications such as eye detection, face detection, facial feature extraction etc. The methods[17][19][20] that claim to have higher accuracy ($e \leq 0.05$) are complex, require some machine learning techniques and computationally expensive. The method[16][18] that claim to be computationally efficient is not highly accurate (for $e \leq 0.05$), which is the prime requirement for efficient eye center localization system. Therefore the aim is to propose a visible spectrum eye center localization method that is accurate (for $e \leq 0.05$) as well as computationally efficient for real time eye center localization in uncontrolled scenarios.

4. OVERVIEW OF PROPOSED APPROACH

A stage wise approach is adopted to perform the eye localization task. The face detection is achieved by Viola Jones boosted cascade classifier[3] in first stage, which performs well in real time to detect faces. The coarse eye regions are derived from the detected face to perform further analysis in second stage. Since the Pixels on the boundary of the iris and pupil have the same intensity, the outlines of iris and pupil can be simplified to circles. The vector field of an image gradient can appropriately be used to detect the center of the circular region. The gradient-based objective function[18] estimates the iris center by searching the maximum value of the function as the iris center. However this method makes redundant searches in the larger eye region and fails to detect eye center in case of prominent eyebrows and eyes obstructed by hair or glasses.

A more efficient approach is proposed here, where a precise coarse eye region is identified using novel geometric eye model derived using face anthropometric parameters, which is further used for eye center localization. Novel range of gradients around the boundary of an iris is proposed that can more precisely contribute to vote for the accurate eye center localization. This approach is faster and eliminates redundant searches. These gradients are then combined with a weight map to detect iris center in an effective and accurate way.

5. PROPOSED GEOMETRIC EYE MODEL AND EDGE GRADIENTS

Face anthropometry parameters are used to identify certain landmark points to detect facial features from the face region. A novel geometric eye model based on these parameters is proposed to identify coarse eye region related to the size of detected faces. Precise eye localization can then be achieved from the coarse eye region derived according to this parameter.

5.1 Coarse Eye Region

According to prendergast[22], the eyes usually measures one-fifth the width and one third the height of the face. This is shown in figure 1 using BioID database. Here w and H are the width and height of the face region detected by Viola Jones[3].



Figure 1 : Face anthropometric measures on BioID database.

Figure 2 shows the proposed geometric eye model to improve the coarse eye region detection. Here a is the width of the coarse eye region, which is approximately $1/5^{\text{th}}$ of the relative face width w , r is the radius of the iris, which is roughly equal to $a/6$. X is the region above and below the iris area, which is estimated to be combinly equal to $a/2$ (i.e $2X=a/2$). Thus the width and the height of the coarse eye region is derived as $w/5$ and $w/6$ ($w/6 = a/2+a/3$) respectively. The derived coarse eye region as per this model is shown in figure 6(a).

$$a = w/5, \quad 2r = a/3, \quad x = a/4, \quad l = d/w \quad (1)$$

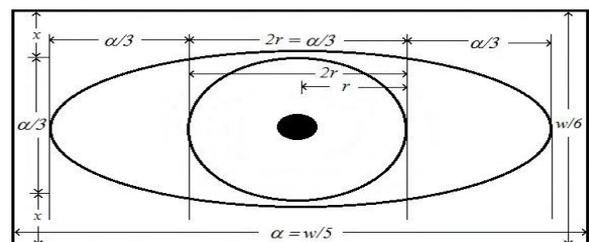


Figure 2 : The proposed Geometric eye model

λ is the ratio of the Euclidian distance d between the two eye centers and face width w as shown in figure 3. Here $0.3 \leq \lambda \leq 0.5$ is derived empirically.

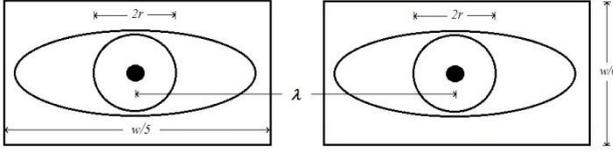


Figure 3 : The ratio (λ) of the Euclidian distance d between two eye center and face width w .

5.2 Eye Center Localization

The coarse eye region image is converted to gray-scale and blurred slightly to remove the noise and reflection. The gradient-based approach[18] is adopted to estimate the iris center, which is represented in figure 4, with the iris and sclera as grey circle and white background. Here, let x_i be an element along the edge of the circle, g_i be the normalized gradient vector at pixel x_i and d_i be the unit vector from c to x_i . If c is the real center, the displacement vector d_i and the normalized vector g_i should have the same orientation and their dot product achieves the max value compared to other locations of c .

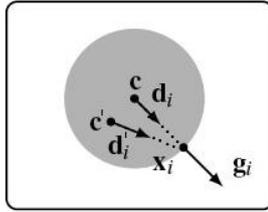


Figure - 4 : The dot product between unit vector d_i and g_i achieves maximum value for center c .

For an image R , with pixel position $x_i, i \in \{1, \dots, N\}$ that has a gradient field g_i , the optimal center c^* of a circular object can be calculated by:

$$c^* = \underset{c}{\operatorname{argmax}} \left\{ \frac{1}{N} \sum_{x_i \in R} (d_i^T g_i)^2 \right\} \quad (2)$$

However, here the votes are generated by gradients for every possible center. The image gradients as per eq (2) are shown in figure 6(b). For an eye image, pixels with smallest grey level are most likely to be pupil. Hence the grey levels of the pixels are inverted, which can be further used in pupil pixel determination. The inverted image is shown in figure 6(c). In the gradient-based approach[18], every gradient g_i in the eye region contribute for the possible optimal center c . Therefore, the gradients away from iris center may also degrade the result. A novel range for pixels is derived that contain the gradients that have the highest possibility to contribute for votes for the accurate eye center localization. This range is defined as β and shown in figure 5. The range β around the boundary of an iris is indicated as the upper and lower bound as r_1 and r_2 from the pupil center within which the gradients are allowed to vote. It is obtained as following :

$$b = 4 \times \int_0^{\beta} \int_0^{\beta} \frac{\rho^{1/2}}{2} r_1^2 d\theta - \int_0^{\beta} \int_0^{\beta} \frac{\rho^{1/2}}{2} r_2^2 d\theta$$

$$b = 2 \times \int_0^{\beta} r_1^2 \rho^{1/2} d\theta - \int_0^{\beta} r_2^2 \rho^{1/2} d\theta$$

$$b = 2 \times \int_0^{\beta} r_1^2 \times \frac{\rho^{1/2}}{2} d\theta - \int_0^{\beta} r_2^2 \times \frac{\rho^{1/2}}{2} d\theta$$

$$b = \int_0^{\beta} r_1^2 - r_2^2 \rho \quad (3)$$

The optimum value for β is obtained empirically by setting the upper limit r_1 and lower limit r_2 to 37 and 27 respectively. The gradients, which are outside the expected upper and lower bound of an iris radius, are then discarded based on these parameters and the gradients in the remaining eye region including the eyebrow part are set to zero and do not contribute to the result. Votes are then generated from the gradients within specified range to identify the most probable center. Figure 6(d) shows the gradients that fall within the range of β as per eq. (3). Figure 6(e) shows the votes that are generated by the gradients that fall within the range of β .

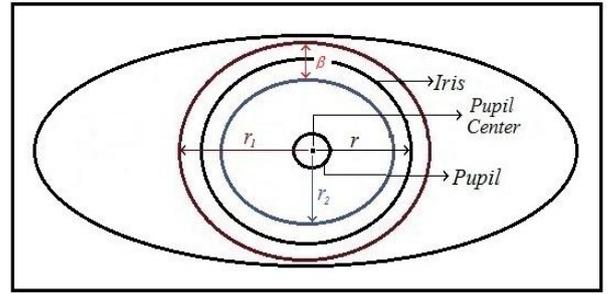


Figure - 5 : The range β around the boundary of an iris within which the gradients are allowed to vote.

This range parameter is applied to both left and right eye images as β_l and β_r , and derived as following.

$$b_l = \int_0^{\beta_l} r_{l1}^2 - r_{l2}^2 \rho \quad \text{and} \quad b_r = \int_0^{\beta_r} r_{r1}^2 - r_{r2}^2 \rho \quad (5)$$

where r_{l1}, r_{l2} and r_{r1}, r_{r2} are the radiuses of the upper bound and lower bound of an iris of the left and right eye respectively. Additionally, the prior knowledge that the pupil centers are roughly λ distance apart is enforced to enhance the accuracy. The pupil is circular and the gradients only in the certain range of β_l and β_r near the radius of an iris and the most common distance of λ are included for calculation of eye centers. Hence, our most probable search region δ that includes β_l and β_r is defined as:

$$d = b_l \oplus b_r \quad | \text{Within } \delta \quad (6)$$

Again, as shown in figure 6(e), if all the votes generated by gradients within the range β is figured as possible candidate center c , the calculation will cost considerable time. A weight map is therefore applied to reduce the number of possible candidate pixel c , defined as w_c where, most of the pixels except the few center pixels turns out to be zero and require no calculation as shown in figure 6(f). This will speedup the task significantly and a pixel with highest value is quickly and accurately identified as an iris center. Finally, the range of pixels obtained from eq (6), integrated into the objective function of eq (2), leads to the ultimate formation expressed in eq (7).

$$c^* = \underset{c}{\operatorname{argmax}} \left\{ \frac{1}{N} \sum_{x_i \in d} w_c (d_i^T g_i)^2 \right\} \quad (7)$$

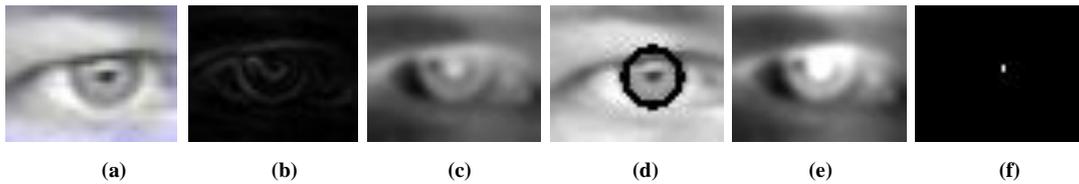


Figure 6 : (a) Coarse eye region as per geometric eye model (b) Image gradients (c) Inverted eye image (d) Gradient contribution within the range of β as per the proposed method. (e) The votes generated by the edge gradients (f) A weight map applied to speedup the calculation.

6. RESULTS AND ANALYSIS

The normalized error proposed by Jesorsky [8] is adopted to calculate the performance of the iris center detection algorithm as per eq (9).

$$e \in \frac{1}{d} \max(e_l, e_r) \quad (9)$$

where e_l and e_r are the Euclidean distance between the estimated iris center locations, and d is the Euclidean distance between the exact left and right iris center locations.

The BioID [21] database is chosen to verify the testing results as it is considered to be a challenging database for eye center localization for its real world situations with variation of scales, expressions, poses, illumination and background condition. The images are provided with the ground truth values, which is used to verify the testing results. However, the images with undetected faces and closed eyes are not considered for the accuracy calculation, as other state of the art methods did. The coarse eye region derived from the proposed geometric eye model turns out to be quite right for BioID database [21] and produce very reliable bounding boxes around the user's eyes as shown in figure 7. It reduces the search space significantly while effectively removing spurious eye detections.

Figure 8 shows accurate eye center localization on BioID database [21] based on proposed geometric model and edge gradients. Table 1 shows the comparison of performance for Normalize Error Measurements for $e \in \{0.05, 0.10, 0.15, 0.20, 0.25\}$ of the proposed method with existing published results of some state of the art methods. Various error values are derived from the respective authors graph where it is not mentioned explicitly. Based on the performance of each method in table1, a rank is derived chronologically for each error value. The comparison of the derived ranks and average rank of each method according to its performance is shown in table 2.

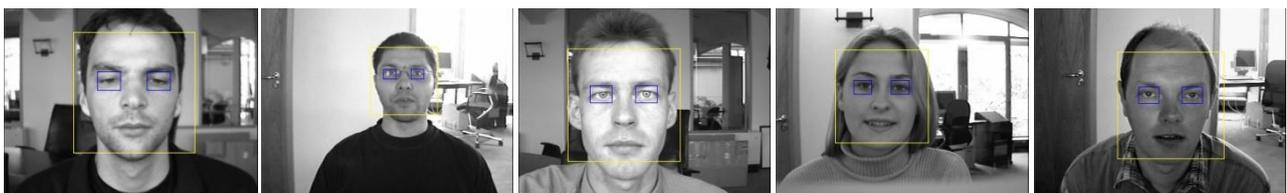


Figure 7 : Accurate coarse eye region identification on BioID database[21] based on proposed geometric model.

6.1 Performance Analysis Based on Normalized Error Value

Eye center localization (for $e \leq 0.05$):

1. The method proposed by Yang[17] yields the best performance in this category with multiple scale sparse dictionary technique which is complex and highly iterative.
2. The proposed method yields a 2nd best performance of 86.51% without high complexity, iteration, clustering or classifier.

Eye center localization ($e \leq 0.10, e \leq 0.15, e \leq 0.20, \text{ and } e \leq 0.25$):

1. Yi[19] provides the best performance for $e \leq 0.10, e \leq 0.15, e \leq 0.20$ and $e \leq 0.25$ with Extending LBP (ELBP) feature set, Stacking classifiers in multiple scales and P-Cascade.
2. The main objective of the proposed method is to improve the accuracy for $e \leq 0.05$ with less computational complexity, which is the desired characteristics for real time eye center localization system however; the proposed method performs well for higher error values with less complexity.

6.2 Performance Analysis based on the Rank

For various methods, a rank is derived according to its performances in table 1. An average rank of individual method is derived based on overall performance of that method. A comparison based on the rank is shown in Table 2.

1. There is no single method that performs superior for all error values.
2. The Method proposed by Yi[19] achieves 1st average rank with overall best performance. However this method has reasonably high complexity and the best performance for $e \leq 0.05$ is not achieved.
3. The proposed method is the 2nd best in accuracy for $e \leq 0.05$ with much less complexity, which is the main objective of the proposed work.

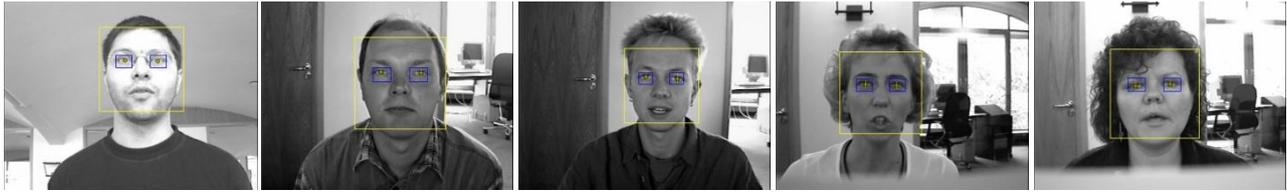


Figure 8 : Accurate eye center localization on BioID database[21] based on proposed geometric model and edge gradients.

Table 1 : Performance Comparison of the proposed methods using Normalized Error Measurement using the BioID database.

Method	$e \leq 0.05$	$e \leq 0.10$	$e \leq 0.15$	$e \leq 0.20$	$e \leq 0.25$
Jesorsky[8]	(38.5%)	(79.0%)	(84.5%)	(87.0%)	91.8%
Cristinacce [11]	(57.0%)	96.0%	(96.5%)	(97.0%)	(97.1%)
Hamouz[12]	(59.0%)	(75.0%)	(80.8%)	(87.6%)	(91.0%)
Niu [13]	(75.0%)	93.0%	(95.8%)	(96.4%)	(97.0%)
Turkan [14]	(18.6%)	73.7%	(94.2%)	(98.7%)	99.6%
Valenti [15]	84.1%	90.9%	(93.8%)	(97.0%)	98.5%
Asadifard [16]*	47.0%	86.0%	89.0%	93.0%	96.0%
Yang[17]	89.60%	95.50%	(97.2%)	(98.5%)	99.10%
Fabian[18] *	82.5%	93.4%	95.2%	96.4%	98.0%
Yi[19]	86.5%	99.1%	(99.3%)	(99.5%)	(99.6%)
Ren[20]	77.08%	92.25%	(95.7%)	(97.0%)	98.99%
Proposed Method *	86.51%	89.90%	90.89%	94.49%	97.74%

() Indicates the values derived from respective author's graph.
*The methods that don't involve any learning or model scheme.

Table 2: Comparison of ranks of each method according to its performance in Table 1.

Method	$e \leq 0.05$	$e \leq 0.10$	$e \leq 0.15$	$e \leq 0.20$	$e \leq 0.25$	Avg. Rank
Jesorsky[8]	11	10	11	12	10	10.8
Cristinacce[11]	9	2	3	4	7	5
Hamouz[12]	8	11	12	11	11	10.6
Niu [13]	7	5	4	7	8	6.2
Turkan [14]	12	12	7	2	1	6.8
Valenti [15]	4	7	8	5	4	5.6
Asadifard [16]	10	9	10	10	9	9.6
Yang[17]	1	3	2	3	2	2.2
Fabian[18]	5	4	6	8	5	5.6
Yi[19]	3	1	1	1	1	1.4
Ren[20]	6	6	5	6	3	5.2
Proposed Method	2	8	9	9	6	6.8

7. CONCLUSION

A novel geometric eye model to identify coarse eye region based on refined anthropometric parameters is proposed, which eliminates problematic eye region and redundant

searches during eye center estimates. Inside this coarse eye region a novel range of parameters around the boundary of an iris are identified within which the gradients are allowed to vote. These will speedup the task, improves the accuracy, and reduces the complexity for eye center localization significantly. The proposed approach provides high accuracy for $e \leq 0.05$, which is the desired objective for the proposed method. For other error values the proposed method performs well with less complexity. Considering the accuracy and efficiency of the proposed method, it can be conveniently and appropriately applied for eye center localization in real time applications.

8. REFERENCES

- [1] B. Froba and C. Kublbeck, "Robust face detection at video frame rate based on edge orientation feature," IEEE Conference on Automatic Face Gesture Recognition, pp. 342–347, May 2002.
- [2] Z. Liu and Y. Wang, "Face detection and tracking in video using dynamic programming," IEEE Proceedings ICIP Vol1, pp. 53–56, 2000.
- [3] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features" IEEE Conference on Computer Vision and Pattern Recognition pp. 511–518, 2001.
- [4] A. Yuille, P. Hallinan and D. Cohen, "Feature extraction from faces using deformable templates," International Journal of Computer Vision, 8:2, 99–111, 1992.
- [5] K. Peng, L. Chen, S. Ruan and G. Kukharev, "A robust algorithm for eye detection on gray intensity face without spectacles." Journal of Computer Science and Technology, Vol. 5, No. 3, pp. 127-132, Oct 2005.
- [6] I. Fasel, B. Fortenberry and J. Movellan, "A generative framework for real time object detection and classification," Computer Vision and Image Understanding, 98(1), 182–210, 2005
- [7] DW. Hansen and JP. Hansen, "Robustifying eye interaction," IEEE International Conference on Computer Vision and Pattern Recognition Workshop pp. 152–159, June 2006
- [8] O. Jesorsky, K. Kirchberg and R. Frischholz, "Robust face detection using the Hausdorff distance" In Proceedings of the 3rd AVBPA, LNCS, Springer pages 90–95, 2001.
- [9] K N Jariwala, U D Dalal, Efficient Performance Evaluation for Robust Eye Localization System, International Journal of Computer Science Trends and Technology (IJCTST), Volume 3 Issue 1, Jan-Feb 2015.
- [10] K N Jariwala and U D Dalal, Performance Analysis of Eye Localization Methods for Real Time Vision Interface using Low Grade Video Camera, International

Journal of Computer Application, Vol. 114 - Number 2, March 2015.

- [11] D. Cristinacce, T. Cootes and I. Scott, "A multistage approach to facial feature detection," In Proceedings of British Machine Vision Conference, Vol. 1, pp 277–286, 2004.
- [12] M. Hamouz, J. Kittler, JK. Kamarainen, P. Paalanen, H. Kalviainen and J. Matas. "Feature-based affine invariant localization of faces," IEEE Transaction on Pattern Analysis and Machine Intelligence, 27(9):1490–1495, 2005.
- [13] Z. Niu, S. Shan, S. Yan, X. Chen and W. Gao, "2D Cascaded AdaBoost for Eye Localization," Proceedings Of International Conference on Pattern Recognition, 2006.
- [14] M. T̄urkan, M. Pard'as and A. Cetin, "Human eye localization using edge projections," Proceedings of the VISAPP, pp. 410–415, 2007.
- [15] R. Valenti and T. Gevers, "Accurate eye center location and tracking using isophote curvature," Proceedings of the CVPR, pages 1–8, 2008.
- [16] M. Asadifard and J. Shanbezadeh, "Automatic adaptive center of pupil detection using face detection and CDF analysis," In Proceedings of the IMECS, volume I, pp. 130–133, 2010.
- [17] F. Yang, J. Huang, P. Yang and D. Metaxas, "Eye Localization through Multiscale Sparse Dictionaries," IEEE International Conference on Automatic Face and Gesture Recognition, FG 2011.
- [18] F. Timm and E. Barth, "Accurate eye centre localisation by means of gradients," in Proceedings of VISAPP, pp. 125–130, March 2011.
- [19] D. Yi, Z. Lei and S. Z. Li, "A robust eye localization method for low quality face images," in Proceedings of the International Joint Conference on Biometrics, pp. 1–6, 2011.
- [20] Y. Ren, S. Wang, B. Hou and J. Ma, "A Novel Eye Localization Method With Rotation Invariance," IEEE Transactions On Image Processing, Vol. 23, No. 1, 2014.
- [21] www.bioid.com/About/BioID-Face-Database
- [22] P. Prendergast, "Facial Proportions," Advanced Facial Surgical rejuvenation, pp 15-22, Springer, 2012.