Computer & Robotics

A Novel Eye Gaze Estimation Method Using Ant Colony Optimizer

Mina Etehadi Abari

Department of Computer Engineering, Amirkabir University of Technology, Tehran, Iran

Received 26 March 2018; Revised 17 January 2019; Accepted 18 February 2020; Available online 11 May 2020

Abstract

This paper addresses the eye gaze estimation problem in low-resolution images, using the low-cost camera in order to eliminate problems caused by infrared high-resolution imaging such as needing an expensive camera, complex setup, special light sources, and being limited in lab research environments. In the proposed method, the human face is detected with Ant Colony Optimization (ACO) algorithm, and then the Kirsch compass mask is utilized to detect the position of humans' eyes. For iris detection, a novel strategy based on ACO algorithm, which has been rarely used before, is applied. The pupil is recognized by morphological processing. Finally, the extracted features, obtained from the radius and position of the irises of the pupils, are given to the Support Vector Machine (SVM) classifier to detect the gaze pointing. In order to receive assurance of the reliability and superiority of the newly designed ACO algorithm, some other metaheuristic algorithms such as (GA, PSO, and BBO) are implemented and evaluated. Additionally, a novel dataset, comprising 700 images gazing at seven different major orientations, is created in this research. The extensive experiments are performed on three various datasets, including Eye-Chimera with 92.55% accuracy, BIOID dataset with 96% accuracy, and the newly constructed dataset with 90.71% accuracy. The suggested method outperformed the state of the art gaze estimation methods in terms of the robustness and accuracy.

Keywords: *Eye-Gaze Estimation, Ant Colony Algorithm, Low-Resolution Image, Kirsch Filter, 2D image.*

1. Introduction

The process of estimating the position that the user is looking at is called eye-gaze estimation and the system which can achieve this aim is known as an eye-gaze tracker.

Intelligent eye-gaze recognition systems can be utilized in the wide array of applications and systems, which have several advantages. The first advantage is that they are beneficial for the disabled, who are bereft of movement. Giving them an opportunity to work, socialize, entertain in their lives, and be independent of others [1, 3, 24]. The second one is that they are practical for computer-novice users. In fact, instead of using a mouse or other input devices, they can easily undertake pointing and selecting tasks by employing the eye-gaze tracker [4]. The third advantage is that they are effective to increase the security of the existing access control systems. Indeed, the aggregation of a user's face and gaze information can play an effective role in this issue [4, 5]. The fourth advantage is in computer games in order to improve gaming excitement. Actually, such systems can be used to integrate gaze input from eye tracker [7, 23]. The fifth advantage is that they are required in the cockpit of combat aircraft instead of joystick-

* Corresponding author. Email: minaetehadi@aut.ac.ir

based pointing. Thus, pilots have the opportunity to dedicate their full attention to controlling the aircraft carefully [4, 15]. The sixth one is they are helpful in advertisement marketing in order to realize customers' interest [5, 7].

Considering the above-mentioned advantages, the importance of this issue motivates us to undertake research on designing a novel method based on the Ant Colony Optimization (ACO) algorithm.

2. Literature Review

The related work on possible solutions for eye-gaze recognition is extensive, spanning more than two decades of research and has observably intensified due to its widespread utilization [1, 5, 12]. However, the first research on eye gaze pointing dates back to the 18th century, in order to explore different eye-gaze styles in a reading task and evaluate psychology and pedagogy of reading [2]. In recent researches, Eye-gaze estimation systems have been divided into two models: appearance-based models and feature-based models. The implementation of the appearance-based model is simple [5]. Appearance-based gaze point recognition systems are bereft of extracting the features explicitly, instead employ the whole image for eye-gaze detection and tracking [5, 12]. However, they cannot accurately localize and track head movements, in spite of extensive attempts in this area such as [10, 11, 23].

The second type, named feature-based gaze estimation method, can be further classified into two groups: 2D-based [13-15] and 3D-based [16, 17] gaze-tracking methods. 3D-based methods are more accurate but need higher costs, complex devices, and longer processing times [17]. In general, the feature-based method contains two phases; first, extracting the proper and discriminated features. Second, object classification based on the obtained feature in the first phase. Wang et al. [18] and Zhang et al. [19] extracted all the edges of iris obtained from the input eye image and just selected the two longest vertical edges as proper features in order to fit the ellipse and find iris center [21, 22]. Our research focuses on 2D-feature-based approach.

Besides the appearance-based and feature-based models, another classification approach for eye-gaze estimation system is based on the type of input images including infrared imaging and visible imaging. The first type requires infrared cameras and infrared light sources to capture the infrared images, whereas the second type employed the usual cameras, which is far less expensive. Due to the fact that in infrared-imaging, invisible infrared light source is applied to obtain the controlled light, it has the ability to decrease the effects of light conditions and generates a severe contrast between the iris and pupil (for example bright-dark eye effect), as well as the reflective features of the pupil and the cornea [12]. Although the infrared imaging-based tracker system has high accuracy, its disadvantages could not be ignored. For instance, it is expensive. Moreover, it is not a reliable method under the disturbance of other infrared sources. Moreover, these systems need a complex setup, special light sources, camera, and monitor. Therefore, they are limited to lab research and difficult to use in natural environments [7, 23]. In addition, the reflection of infrared light sources on glasses makes a lot of problems. Another drawback of the Infrared imaging-based system is that not all users generate the bright-dark effect and it can cause problems for the eye-gaze tracker [5].

In comparison with the infrared-imaging-based systems, visible imaging models can eliminate the above-mentioned shortcomings without the need for the determined infrared devices. Thus, some recently related scholars adopted 2D visible image-based method in this field [5, 15, 23, 24]. In this approach, we utilize 2D visible-imaging with a usual camera as well in order to obliterate the difficulties caused by infrared-imaging-based systems.

3. General Methodology

The General methodology of this novel scheme is illustrated in Fig.1.

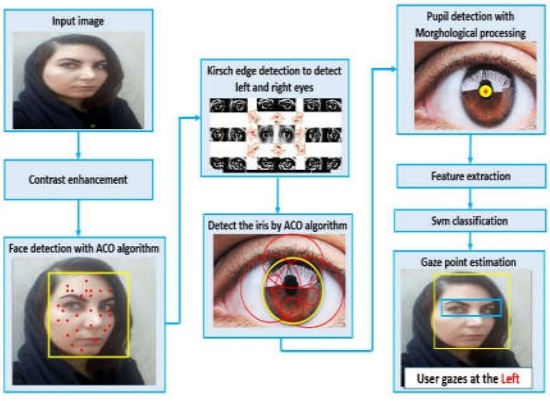

Fig. 1. The methodology of the proposed scheme.

As can be observed in Fig.1, the proposed model contains five main phases as follows:

1. After preprocessing and contrast enhancements of images, the face of the user has been detected with Ant Colony Optimization (ACO) algorithm.
2. After cropping the face from the other part of the input image, the position of the left and right eye is extracted by utilizing the Kirsch compass masks as an edge detector.
3. The novel proposed strategy based on the ACO algorithm is utilized at this step to detect the position of the iris.
4. At this phase, the pupil is recognized with the morphological processing.
5. The ten extracted features, obtained from the radius and position of the irises and pupils, are given to the SVM classifier to estimate the eye- gaze point.

In the next sections, each phase is described comprehensively.

3.1. Face Detection

This paper aims to employ the ACO algorithm to detect the faces of the users. The ACO was suggested by Dorigo et al. for the first time [27]. It is an evolution simulation algorithm and is inspired by the operations of the real ants for food hunting.

All ants individually move from the beginning node to the last node for searching food [27, 28, 29]. The probability of ant k in node i moving to node j at time t is denoted in Equation 1:

$$p_{ij}^{k} = \frac{\tau_{ij}^{\alpha}(t).\eta_{ij}^{\beta}(t)}{\sum_{l\in N_i^k} \tau_{ij}^{\alpha}(t).\eta_{ij}^{\beta}(t)} \quad (1)$$

Where τ is the pheromone trail, η is the inverse of the cost (i, j), which is the cost between node i and j and denotes the shortest path from the node (pixel) i to node j, α and β determine the relative importance of τ to η. N_i^k is the set of nodes (pixels) that remain to be visited by ant k In Eq. (1) if α = 0, no pheromone information is considered. If β = 0, the attractiveness of moves is ignored. The best parameters of the proposed ACO algorithm that are reached in our experiments are as follows: (α=0.6, β=0.4, ρ = 0.3) Moreover, numbers of ants in each iteration and numbers of iterations are customized to 50. Various quantities of these parameters (α, β, ρ, numbers of ants, and numbers of iterations) are tested on 100 face images in our experimentations and the best ones are selected. Other valid scholars such as [28, 29, 31] have used this method to acquire the best quantities of these parameters as well. At t=0, trail (τ) = 0, (η) = 0, and 50 ants are assigned to the pixels accidentally. In fact, at the first iteration, each ant is assigned to one pixel by random and the ants release pheromone trail on the visited pixels (node) which satisfy the illumination criteria and other visited pixels, which are not the face area, are removed from the searching tour of the upcoming ants. At the second iteration, the 50 new generated ants are assigned to the pixels based on the previous acquired trail (τ) and the goodness of moves (η). This process is repeated for 50 iterations and the best iteration which contains more pixels of the face area is selected. Finally, the smallest square, which contains all the desirable ants of the best iteration (ants which satisfy illumination criteria and are located in the face area), is generated and selected as the face area.

When the ants complete the tour successfully for one time, then the amount of pheromone in the link between node *i* and *j* would be modified according to [27, 28]:

$$\tau_{ij}(t+1) = (1-\rho)\,\tau_{ij}(t) + \rho \sum_{k=1}^{K} \nabla\tau_{ij}^{k} \quad (2)$$

$$\nabla\tau_{ij}^{k} = \begin{cases} \frac{Q}{d_{ij}} & \text{If ant k goes from edge (pixel) i to j between t and t+1} \\ 0 & \text{Otherwise} \end{cases} \quad (3)$$

where $\nabla\tau_{ij}^{k}$ is the quantity of pheromone laid on edge (*i,j*), *K* is total numbers of ants, ρ ($0 < \rho < 1$) is the decay parameter and means that the pheromone on the link (*i,j*) evaporates based on this multiplicative modulus. *Q* is a constant and denotes the deposited pheromone quantity on the entire path_{ij}. d_{ij} refers to distance from pixel i to pixel j. The amount of pheromone deposited is inversely proportional to the distance from i to j [27, 31].

The process of face detection using ACO algorithm is illustrated in Algorithm 1. After extracting R, G, B color space from the input image and calculating the size of the input image, the ACO method is adopted. In each tour, the cost function is calculated. At the end of each iteration, the best cost, which is the minimum one, is chosen. In the proposed ACO algorithm, pixels are nodes. It means ants visit pixels (or nodes). If the visited pixel (node) has the required illumination criteria for face detection, the ACO

saves that node; otherwise that node (pixel) is eliminated from the searching tour of ants. The illumination criteria, which is illustrated in Algorithm 1, Step 5, is obtained experimentally. Indeed, the values of R, G, B for skin part is taken from 100 various faces in our experimentation. We calculated these ranges by using each image and enchanting R, G, B values that lie on face part and with some mathematics equations a relationship of R, G, and B values for a skin part in an image is built. Ants should search for pixels (nodes) that have R, G, B values of the face. It should be mentioned that the number of ants in each iteration is 50, and the cost function is (50- number of the ants which are located in the face area.). Then, the best group of nodes (pixels), which have the lowest cost are selected. These selected pixels which have the lowest cost and satisfy the illumination criteria can help us to detect the face. It means the smallest generated square which contains all these pixels is considered as the face area.

Fig. 2 displays an example of the face detection process using the proposed ACO algorithm.

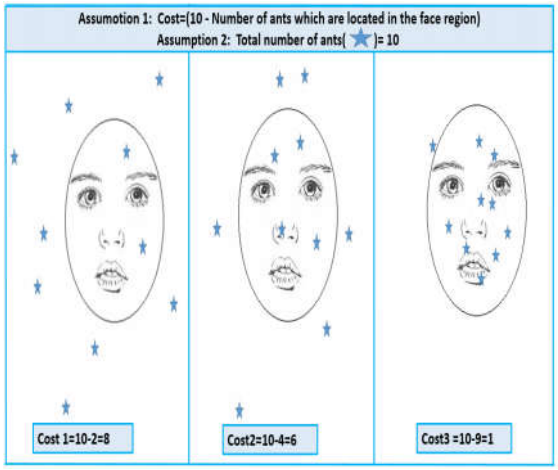

Fig. 2. An example of face detection with the ACO.

Algorithm 1: Face detection using ACO algorithm

Step 1: Extract the R, G, B color space from the input image.

Step 2: Calculate the image's size.

Step 3: Initializing the ACO parameters:

Maximum number of Iteration =50

Number of ants in each iteration=50

Step 4: Start ACO algorithm from 1 to the number of iterations.

Step 5: Calculate the cost function and find the lowest cost.

Cost function= (50- number of ants which are located in the face area.)

Illumination criteria for face detection

(based on our dataset)

1. Red > 95, Green > 40, Blue > 20
2. (Maximum illumination of image) – (Minimum illumination of image) >15
3. Red-Green >20,
4. Red-Blue >30

Step 6: At the end of each iteration, the lowest cost is chosen.

Step 7: At the end of 50 iterations, the smallest square, which contains all the desirable ants of the best iteration (ants which are located in a face area and satisfy illumination criteria), is generated and considered as the face area.

The question which is often raised here is if is it possible to extract the face by finding the biggest region, which satisfies the defined illumination criteria without ACO algorithm? Extensive experiments on our dataset have shown that if we just use the illumination criteria in order to detect faces, it takes high computational time. Conversely, by using the proposed ACO, it takes less computational time and detects the faces with higher speed. Since in our new ACO algorithm, in each iteration, not only the pixels which match the illumination criteria are more likely to be chosen as the correct regions thanks to the released pheromone trail, but also the pixels which do not satisfy the illumination criteria (are not face region) are removed from the searching process of the next iterations. By this wise decision, the speed process of precise convergence to the face area is increased and the running time of face-detection is decreased as well. In fact, by utilizing our ACO algorithm about ~30% of all pixels are optimally analyzed while without ACO all pixels of the given image should be analyzed which is time-consuming. Moreover, dark skin faces, which were hard to be detected by using the illumination criteria, can be detected by using the proposed ACO algorithm thanks to the precise convergence of ants to the face area.

The superiority of the proposed ACO algorithm in comparison with another metaheuristic face detection method based on Genetic Algorithm (GA) [33], is that our strategy is more robust and accurate and detect a wide range of faces with high accuracy rate while [33] could not detect faces with dark skins.

Furthermore, another novel method suggested in 2014 [34], which was based on finding the biggest region that could satisfy the illumination criteria, came down with a

drawback. It needed a long-time process while our proposed ACO method is fast and can detect the faces more quickly.

3.2. Detection of Key Points

After cropping the face from the other part of the input image, the key points of the face should be detected. These key points are the position of the left and right eye and the location of the iris and pupils. Afterward, we find a relation between the features of these key points and estimate the eye gaze pointing.

3.2.1. Eye Recognition

Before explaining the offered method for eye recognition, let overview the structure of the eye for further elucidation [48]. Each eye image can be divided into three regions: high-intensity region (sclera), medium-intensity region (eyelid), and low-intensity region (iris and pupil). The low-intensity region including iris and pupil is surrounded by sclera as higher-intensity regions. In fact, as shown in Fig.3 the eyeball consists of three kinds of the iris boundary: inner iris boundary (pupil–iris), outer iris boundary (iris–sclera), and the eyelid boundary (iris–eyelid).

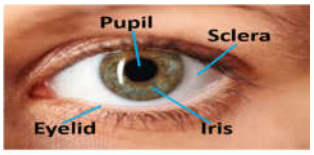

Fig. 3. The eye structure.

In this paper, Kirsch operator is utilized as a mask for edge detection. Kirsch operator [37], discovers maximum edge strength in a few pre-determined directions. It takes a single kernel mask and rotates in ⟦45⟧^(□)increments through all 8 major compass orientations including North, South, East, West, South West, South East, North West, North East [35, 36]. The Kirsch compass masks, including 8 template matrix, is illustrated in Fig.4.

$$M_0 = \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \quad M_1 = \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} \quad M_2 = \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \quad M_3 = \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$$

$$M_4 = \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} \quad M_5 = \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} \quad M_6 = \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} \quad M_7 = \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}$$

Fig. 4. The Kirsch compass masks.

Fig. 5 depicts a sample image in which the Kirsch operator is applied and finally comes with the correct region of the left and right eye as its output. It should be noted that before applying Kirsch operator, the bottom part of the extracted face is eliminated since we are sure that the bottom part of the face including lips and nose are not needed for eye gaze estimation. As a result, Kirsch mask edge detector is applied only on the up half part of the extracted face.

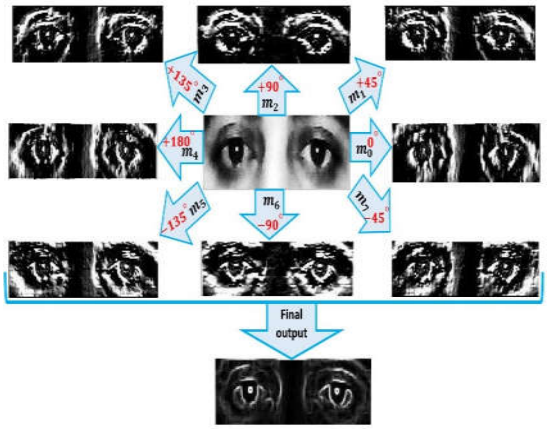

Fig. 5. Using the Kirsch compass masks to approximate the position of the eye.

3.2.2. Iris Detection

The novel tailored ACO algorithm is used to detect the outer iris boundary (iris–sclera). ACO algorithm was introduced in section 3.1 for face detection. Here, the maximum number of iteration is 50 and the number of ants in each population is 100. At each iteration, 10 different circles with random radiuses are generated and thrown on the image. The radius of these circles should be less than the radius of the eyeball. (Eye ball was obtained by using the Kirsch mask in the previous step). We evaluated 100 various iris of different people in our experimentation and we realized that the gray level scale of the iris boundary is bigger than 220 (220 out of 256). In each iteration, ants should search for pixels which have more than 220 gray level and are located on the generated circle boundary. The cost function is (100 – number of ants which are located on the circle boundary and are bigger than 220gray level which is illumination criteria for iris boundary). At the end of each iteration, the lowest cost (the best fitted- circle) is chosen. Consequently, the best-selected circle, which contains more ants and less cost, is the most similar one to the iris. Fig.6 can simplify the process of iris estimation.

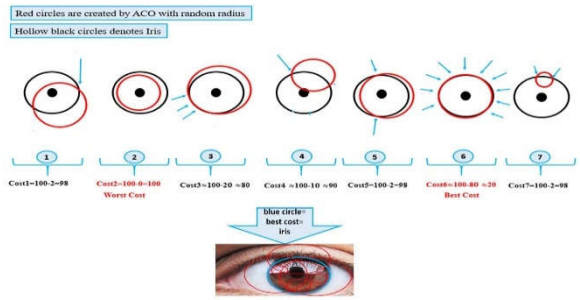

Fig. 6. The process of iris detection using ACO algorithm.

3.2.3. Morphological Processing for Pupil Detection

In order to recognize the pupil, we should realize whether there is a black disk in the input image since the pupil is like a black circular disk. On account of the fact that the pupil is surrounded by many noises such as eyelashes and high light regions in the pupil, the entirety of the pupil is ruined. In order to solve this problem, mathematical morphology is utilized to decrease such noises. The major morphological operations are closing and opening [42]. We refer to [42, 43] in order to introduce these operations. The opening of set X, by structuring element B is:

$$X \circ B = (X \ominus B) \oplus B \tag{4}$$

where $\oplus$, o, and Θ depict dilation, opening, and erosion, respectively. The opening of X by B is simply the erosion of X by B, followed by a dilation by B. The opening operation is based on shape, for instance, only those regions that comprise the structure element can pass the filter. The closing of set X, by structuring element B is [43]:

$$X \bullet B = (X \oplus B) \ominus B \tag{5}$$

In other words, the closing of X by B is simply the dilation of X by B.

In this paper, the opening operation (eq.(4)) is applied to remove eyelashes and the operation closing (eq.(5)) is used to fill the small holes in the pupil. The used mathematical morphology operations are illustrated in Fig.7.

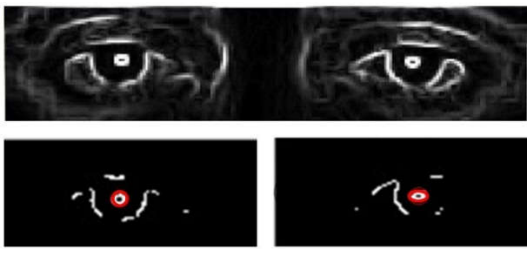

Fig. 7. Mathematical morphology operations for pupil detection (red circle shows pupil).

3.3. Feature Extraction and SVM Classification

For the left eye, 5 features including the radius of iris(r_i), the position of the iris (x_i , y_i), and the position of the pupil (x_p , y_p) are extracted. Similarly, this process is done for the right eye. In order to normalize and extract these parameters, Daugman Rubber Sheet Model [7] is utilized. This practical model is also used to obtain the coordinate of the pupil and iris in some papers such as [7, 8, 24]. After feature extraction, SVM [38] is employed for feature learning and classification task.

4. Experimental Results

The experiments contain two sections. The first section elucidates the face detection experiments while the next one explains the eye-gaze estimation process.

4.1. Experiments on Face Detection

As mentioned earlier, a novel face detection strategy, based on ACO algorithm, is used in the tailored system. To evaluate the performance of the suggested method, a dataset is needed. Thus, a new dataset including 700 face images is constructed. Half of these images are male faces and half of them are female faces. The novel conducted dataset was recorded in the natural illumination condition with a usual camera to evaluate and test face detection in low-resolution.

The PASCAL measure [40] has been employed, as is standard. The detection is true if the overlap of the ground truth annotation and the detection bounding box is more than 50%:

$$a_0 = \frac{\text{are a}(BB_{dt} \cap BB_{gt})}{\text{are a}(BB_{dt} \cup BB_{gt})} > 0.5 \tag{6}$$

Where BB_{gt} is bounding box of the ground truth annotation and BB_{dt} demonstrates the bounding box of the detection system.

In order to be assured of the validity and reliability of the proposed method, different metaheuristic algorithms, are applied for face detection task and the designed ACO algorithm acquires a superior performance ultimately. Indeed, various approaches including genetic algorithm (GA) [33, 41], biogeography-based optimization (BBO) [14], particle swarm optimization (PSO) [9, 20] and ACO are implemented in equal conditions. The obtained outputs

demonstrate that ACO algorithm has high accuracy rate in face detection. Fig.8 can provide more legitimacy and support for this assertion. As can be observed from it, in 35 iterations, the ACO algorithm encounters fewer errors compared to the other scrutinized algorithms. Fig.9 illustrates 2 different samples of the constructed dataset. It shows some face samples which could not be recognized by GA, PSO, and BBO algorithms but the ACO algorithm has the ability to detect these faces precisely.

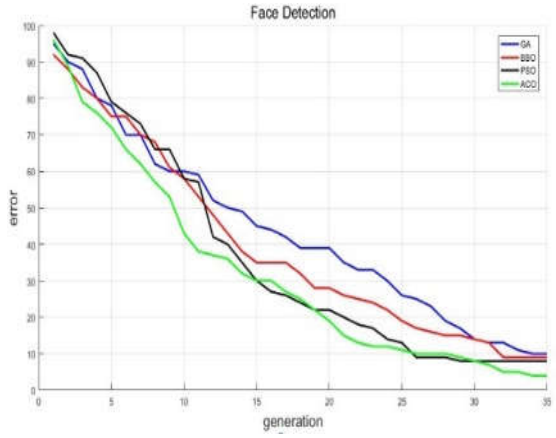

Fig. 8. The results of evaluating different algorithms for face detection.

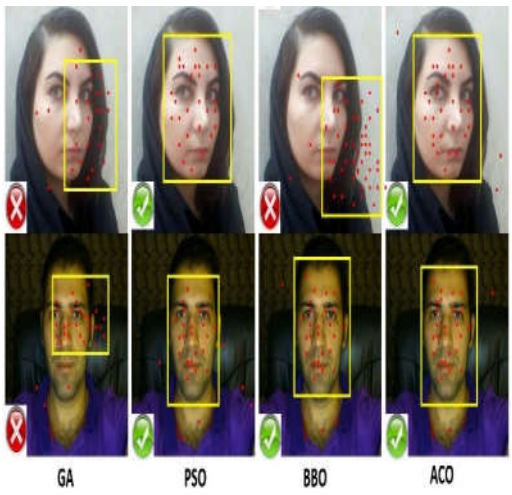

Fig. 9. Two Samples of our constructed dataset using GA, PSO, BBO, and ACO algorithm for face detection.

4.2. Experiments on gaze point estimation

At this part, the experimental results of gaze point recognition are explained. Ten subjects, five males, and five females volunteered to participate in the experiment. Half of the subjects had black or brown eyes and the remained subjects had green or blue eyes. The average age was 28. The newly created dataset comprises 700 images recorded in the natural illumination condition with a usual camera to evaluate and test gaze estimation in low-resolution images (30*15). Subjects gazed at 7 different major orientations. For each gaze direction, 100 different samples are recorded.

Moreover, different metaheuristic algorithms, are applied and the tailored ACO algorithm acquires a superior performance ultimately. The captured outputs prove that the ACO algorithm has high accuracy rate in iris detection. Fig.10 can shed more light on our assertion. In addition, Fig.11 denotes 3 different samples. It shows some eye gaze points which could not be recognized by GA, PSO, and BBO algorithms but are estimated by ACO algorithm exactly.

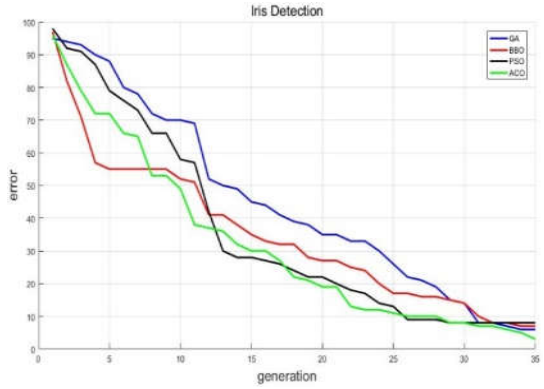

Fig. 10. The results of evaluating different algorithms for iris detection in 35 iterations.

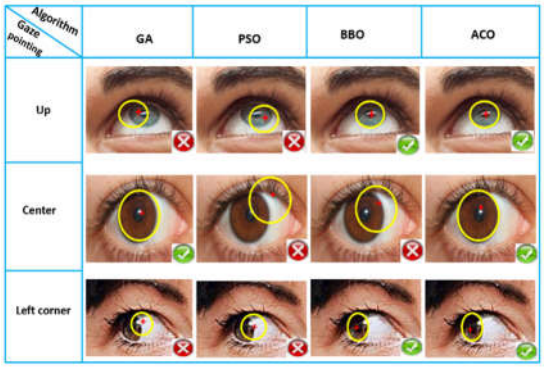

Fig. 11. Three samples of eye gaze estimation with different algorithms in 35 iterations.

Another way to evaluate the suggested method is using the confusion matrix. The 7 various orientations are determined and the gaze estimation points are compared

with the ground truth points. The result is depicted in Table.1. It shows that the proposed method robustly and precisely estimates the orientation that the user is looking. In most items, the gaze error occurs in the y-direction compared to the x-direction since the range of eye movement in the y-direction is lower than that in the x-direction. The average accuracy for all the 7 directions is 90.71% on our created dataset.

Table 1: Physical parameters of the system

Predicted Direction Actual Direction	center	North-east	North-west	East	West	South-east	South-west
Center	96%	0%	1%	0%	3%	0%	0%
North-east	4%	89%	1%	6%	0%	0%	0%
North-West	2%	1%	90%	0%	7%	0%	0%
East	1%	1%	0%	96%	0%	2%	0%
West	0%	0%	4%	0%	96%	0%	0%
South-east	3%	0%	0%	5%	1%	83%	8%
South-West	0%	0%	0%	1%	7%	7%	85%

4.3. Experiments on gaze point estimation

In order to compare the offered method with state of the art, a common dataset is required for test and evaluation. Consequently, eye images of the Eye-Chimera dataset [39] and BIOID dataset [50] are utilized since some previous relevant studies, [9, 25, 32, 39], have used these datasets as well. Additionally, a standard measure is needed for assessment. The accuracy of the iris localization is determined as the largest of the normalized errors for both eyes. The normalized error [32] is calculated using.

$$e = \frac{\max d_{le}}{d_e} \quad (7)$$

Where d_e illustrates the Euclidean distance between left and right iris centers of ground truth eyes and d_{le} represents the Euclidean distance between the recognized center of the iris and ground truth left eyes. The normalized error of the right eye is calculated as the same as left eye. Table.2 and Table.3 depict the accuracy of our method in comparison with previous methods. It is compared to two cases depending on the range of the normalized error. In the range of e <= 0.05, the estimated iris center is exactly located in the pupil area [32] and in the range of e <= 0.1, the estimated iris center is roughly located in the iris area. As is obvious, the suggested method has an acceptable result. Fig.12 displays a sample result of the suggested method.

Table 2: Performance of different methods on BIOID dataset

Different methods using BIOID dataset	Accuracy (e<= 0.1)	Accuracy (e<= 0.05)
Campadelli et al. [26]	85.20%	62.00%
Niu et al. [9]	93.00%	75.00%
Valenti et al. [40]	91.67%	86.00%
Cheung et al. [5]	93.42%	87.21%
Proposed method	96%	90.60%

Table 3: Performance of different methods on Eye-Chimera dataset

Different methods using Eye-Chimera dataset	Accuracy (e<= 0.1)	Accuracy (e<= 0.05)
Valenti et al. (2012) [40]	50.59%	16.10%
Timm et al. (2011) [25]	67.80%	33.60%
Florea et al. (2013) [39]	78.70%	65.30%
Kim et al. (2017) [30]	86.51%	82.34%
Proposed method	92.51%	89.41%

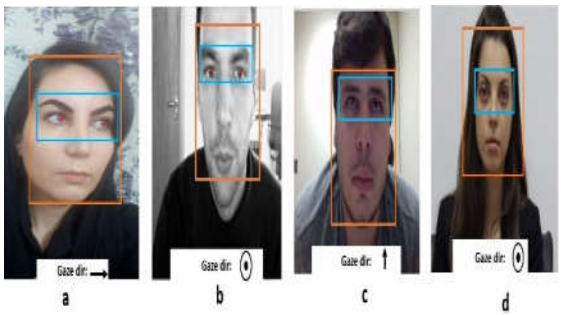

Fig. 12. Experimental results of the proposed method on our constructed dataset, BIOID, and Eye-Chimera dataset.

5. Conclusion

A robust accurate, feature-based model for eye gaze estimation has been constructed in this contribution. Its primary novelty is using a novel designed ACO algorithm, which has not been utilized before. Additionally, a novel dataset, containing 700 images and subjects gazed at 7 different major orientations, is created in this research. Furthermore, it has the ability to detect the eye-gaze in low-resolution images while many previous studies focused on high-resolution images. In addition, 2D visible imaging with the usual camera is employed in this system to eliminate problems caused by using the infrared imaging such as needing an expensive camera, complex setup, special light sources, and being limited in lab research environments.

Extensive experiments are performed on three datasets including Eye-Chimera with 92.55% accuracy, BIOID with 96% accuracy, and the newly created dataset with 90.71% accuracy. Each of these datasets is recorded in different environments with various image resolutions and light conditions. In conclusion, the suggested method outperformed the state of the art methods in terms of robustness and precision.

References

[1] Kim, B. C.; Ko, D.; Jang, U.; Han, H.; Lee, E. C., 3D Gaze tracking by combining eye-and facial-gaze vectors. The Journal of Supercomputing, vol. 73, no. 7, pp. 3038-3052 (2017).

[2] Stanovich, K. E., "Concepts in developmental theories of reading skill: Cognitive resources, automaticity, and modularity" Developmental review, vol. 10, no. 1, pp. 72-100 (1990).

[3] Armato, A.; Lanatà, A.; Scilingo, E. P., "Comparative study on photometric normalization algorithms for an innovative, robust and real-time eye gaze tracker", Journal of real-time image processing, vol. 8, no. 1, pp. 21-33 (2013).

[4] Biswas, P.; Langdon, P., "Multimodal intelligent eye-gaze tracking system", International Journal of Human-Computer Interaction, vol. 31, no. 4, pp. 277-294 (2015).

[5] Chen, J.; Ji, Q., "A Probabilistic Approach to Online Eye Gaze Tracking Without Explicit Personal Calibration", IEEE Transactions On Image Processing, vol. 24, no. 3, pp. 1076-1086 (2015).

[6] Maheswari, S. U.; Anbalagan, P.; Priya, T., "Efficient iris recognition through improvement in iris segmentation algorithm", International Journal on Graphics, Vision and Image Processing, vol. 8, no. 2, pp. 29-35 (2008).

[7] Jagadeesan, A.; Thillaikkarasi, T.; Duraiswamy, K., "Cryptographic key generation from multiple biometric modalities: Fusing minutiae with iris feature", Int. J. Comput. Appl, vol. 2, no. 6, pp. 16-26 (2010).

[8] Krisshna, N. A.; Deepak, V. K.; Manikantan, K.; Ramachandran, S., "Face recognition using transform domain feature extraction and PSO-based feature selection", Applied Soft Computing, vol. 22, pp. 141-161 (2014).

[9] Tan, K. H.; Kriegman, D. J.; Ahuja, N., "Appearance-based eye gaze estimation", In Applications of Computer Vision, (WACV 2002). Proceedings. Sixth IEEE Workshop on pp. 191-195 IEEE (2002).

[10] Lu, F.; Okabe, T.; Sugano, Y.; Sato, Y., "A head pose-free approach for appearance-based gaze estimation", In BMVC, pp. 1-11 (2011) .

[11] Sugano, Y.; Matsushita, Y.; Sato, Y.; Koike, H., "An incremental learning method for unconstrained gaze estimation", In European Conference on Computer Vision, Springer, Berlin, Heidelberg, pp. 656-667 (2008).

[12] Kim, H.; Lee, S. H.; Sohn, M. K.; Kim, D. J., "Illumination invariant head pose estimation using random forests classifier and binary pattern run length matrix", Human-centric Computing and Information Sciences, vol. 4, no. 1 (2014).

[13] Simon, D., "Biogeography-based optimization", IEEE transactions on evolutionary computation, vol. 12, no. 6, pp. 702-713 (2008).

[14] Cho, D. C.; Yap, W. S.; Lee, H.; Lee, I.; Kim, W. Y., "Long range eye gaze tracking system for a large screen", IEEE Transactions on Consumer Electronics, vol. 58, no. 4 (2012).

[15] Bostanci, E.; Kanwal, N.; Clark, A. F., "Augmented reality applications for cultural heritage using Kinect", Human-centric Computing and Information Sciences, vol. 5, no. 1 (2015).

[16] Wang, J. G.; Sung, E., "Study on eye gaze estimation", IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 32, no. 3, pp. 332-350 (2002).

[17] Wang, J. G.; Sung, E.; Venkateswarlu, R., "Estimating the eye gaze from one eye", Computer Vision and Image Understanding, vol. 98, no. 1, pp. 83-103 (2005).

[18] Zhang, W.; Zhang, T. N.; Chang, S. J., "Eye gaze estimation from the elliptical features of one iris", Optical Engineering, vol. 50, no. 4 (2011).

[19] Eberhart, R.; Kennedy, J., "A new optimizer using particle swarm theory", In Micro Machine and Human Science, 1995. MHS'95, Proceedings of the Sixth International Symposium on (pp. 39-43). IEEE (1995).

[20] Valenti, R.; Gevers, T., "Accurate eye center location through invariant isocentric patterns", IEEE transactions on pattern analysis and machine intelligence, vol. 34, no. 9, pp. 1785-1798 (2012).

[21] Valenti, R.; Gevers, T., "Accurate eye center location and tracking using isophote curvature", In Computer Vision and Pattern Recognition, CVPR 2008. IEEE Conference on pp. 1-8 (2008).

[22] Cho, D. C.; Kim, W. Y., "Long-range gaze tracking system for large movements", IEEE Transactions on Biomedical Engineering, vol. 60, no. 12, pp. 3432-3440 (2013).

[23] Baek, S. J.; Choi, K. A.; Ma, C.; Kim, Y. H.; Ko, S. J., "Eyeball model-based iris center localization for visible image-based eye-gaze tracking systems", IEEE Transactions on Consumer Electronics, vol. 59, no. 2, pp. 415-421 (2013).

[24] Timm, F.; Barth, E., "Accurate Eye Centre Localization by Means of Gradients" Visapp, vol. 11, pp. 125-130 (2011).

[25] Campadelli, P.; Lanzarotti, R.; Lipori, G., "Precise Eye Localization through a General-to-specific Model Definition", In BMVC vol. 1, pp. 187-196 (2006).

[26] Dorigo, M.; Birattari, M., "Ant colony optimization", In Encyclopedia of machine learning, pp. 36-39 (2011).

[27] Prakasam, A.; Savarimuthu, N., "Metaheuristic algorithms and probabilistic behaviour: a comprehensive analysis of Ant Colony Optimization and its variants", Artificial Intelligence Review, vol. 45, no. 1, pp. 97-130 (2016).

[28] Hao, Z.; Zhang, X.; Yu, P.; Li, H., "Video object tracing based on particle filter with ant colony optimization", In Advanced Computer Control (ICACC), 2010 2nd International Conference on vol. 3, pp. 232-236. IEEE (2010).

[29] Kim, H. I.; Kim, J. B.; Park, R. H., "Efficient and Fast Iris Localization Using Binary Radial Gradient Features for Human–Computer Interaction", International Journal of Pattern Recognition and Artificial Intelligence, vol. 31, no. 11, (2017).

[30] Chen, B.; Chen, L.; Chen, Y., "Efficient ant colony optimization for image feature selection", Signal processing, vol. 93, no. 6, pp. 1566-1576 (2013).

[31] Valenti, R.; Gevers, T., "Accurate eye center location through invariant isocentric patterns", IEEE transactions on pattern analysis and machine intelligence, vol. 34, no. 9, pp.1785-1798 (2012).

[32] Srivastava, D. K.; Budhraja, T., "An Effective Model for Face Detection Using R, G, B Color Segmentation with Genetic Algorithm", In Proceedings of First International Conference on Information and Communication Technology for Intelligent Systems, vol. 2, pp. 47-55 (2016).

[33] Kumar, P.; Shashidhara, M., "Skin color segmentation for detecting human face region in image", In Communications and Signal

Processing (ICCSP), 2014 International Conference on pp. 001-005. IEEE (2014).

[34] Rivera, A. R.; Castillo, J. R.; Chae, O. O., "Local directional number pattern for face analysis: Face and expression recognition", IEEE transactions on image processing, vol. 22, no. 5, pp. 1740-1752 (2013).

[35] Zhang, T.; Tang, Y. Y.; Fang, B.; Shang, Z.; Liu, X., "Face recognition under varying illumination using gradientfaces", IEEE Transactions on Image Processing, vol. 18, no. 11, pp. 2599-2606 (2009).

[36] Kirsch, R. A., "Computer determination of the constituent structure of biological images", Computers and biomedical research, vol. 4, no. 3, pp. 315-328 (1971).

[37] Cortes, C.; Vapnik, V., "Support-vector networks", Machine learning, vol. 20, no. 3, pp. 273-297 (1995).

[38] Florea, L.; Florea, C.; Vrânceanu, R.; Vertan, C., "Can Your Eyes Tell Me How You Think? A Gaze Directed Estimation of the Mental Activity", In BMVC (2013).

[39] Valenti, R.; Sebe, N.; Gevers, T., "Combining head pose and eye location information for gaze estimation", IEEE Transactions on Image Processing, vol. 21, no. 2, pp. 802-815 (2012).

[40] Goldberg, D. E., "Genetic Algorithms in Search, Optimization, and Machine Language", Addison-Wesley: Reading, UK (1989).

[41] Luengo-Oroz, M. A.; Faure, E.; Angulo, J., "Robust iris segmentation on uncalibrated noisy images using mathematical morphology", Image and Vision Computing, vol. 28, no. 2, pp. 278-284 (2010).

[42] Haralick, R. M.; Sternberg, S. R.; Zhuang, X., "Image analysis using mathematical morphology", IEEE transactions on pattern analysis and machine intelligence, vol. 4, pp. 532-550 (1987).