# Iris Center Localization Using Energy Map With Image Inpaint Technology and Post-Processing Correction

**LIHONG DAI**[1,3]**, JINGUO LIU**[1,2]**, (Senior Member, IEEE),**
**ZHAOJIE JU**[1,2,4]**, (Senior Member, IEEE),**
**AND YANG GAO**[5]**, (Senior Member, IEEE)**

[1]State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China
[2]Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, China
[3]University of Chinese Academy of Sciences, Beijing 100049, China
[4]School of Computing, University of Portsmouth, Portsmouth P01 3HE, U.K.
[5]Space Technology for Autonomous and Robotic Systems Laboratory (STAR LAB), Surrey Space Centre, University of Surrey, Guildford GU2 7XH, U.K.

Corresponding author: Jinguo Liu (liujinguo@sia.cn)

**ABSTRACT** Iris center localization is the basis of iris biometrics, face recognition and gaze tracking. However, individual differences, changes in facial expression, varying light conditions, occlusion, and so on, all bring great challenges to accurately localize the iris center. In order to improve localization accuracy in low-quality images and meet the need of efficiency in practical applications, a novel method of iris center localization is proposed in this paper using energy map synthesis based on image gradient, image inpaint technology, and post-processing correction. The image inpaint technology is firstly adopted to inhibit the effect of some specular reflection. Then the energy maps based on image gradient and eye ROI (Region Of Interest) midpoint are synthesized to significantly improve the localization accuracy. In the end, post-processing correction is carried out to eliminate influence of the closed eye and other large derivations to further improve the localization accuracy. The algorithm is verified on the challenging BioID database, Talking Face Video database and the MUCT face database. The result shows the localization accuracy has outperformed the state-of-the-art unsupervised methods on the three databases, and it is suitable for real-time applications.

**INDEX TERMS** Iris center localization, image gradient, image inpaint, energy map synthesis, post-processing correction.

## I. INTRODUCTION
### A. SIGNIFICANCE AND CHALLENGE OF IRIS CENTER LOCALIZATION

Iris center localization is the basis of iris biometrics, face recognition and gaze tracking, whose accuracy mainly depends on that of iris center localization. However, individual differences, changes in facial expression, varying light conditions, occlusions, and so on, have a significant impact on accurate iris center localization, as follows.

1) People from different countries and races have different iris colors and sizes; even for the people of the same race, the iris of different individual is also different.
2) Facial expressions also have an effect on iris center localization. For example, crying or laughing will lead to different degrees of eye deformation. Especially when the eyes are closed, it is difficult to locate the iris center.
3) Different light intensities affect pupil sizes. When exposed to strong light, the pupils become smaller, whereas when the light is darker, the pupils become larger. The most importance is that light has a great

The associate editor coordinating the review of this manuscript and approving it for publication was P. K. Gupta.

influence on the image quality. Too strong or too weak light will make the image quality decline, thus affecting the iris center localization. In addition, the specular reflection of light often causes bright spots and even covers the iris area, which seriously affects the localization accuracy of the iris center.

4) The occlusion of the eye area by hair and glasses also makes it difficult to locate the iris center. Moreover, hair, eyebrows, and black glasses are darker, which also bring challenges to the methods of iris center localization based on image gray features.

However, most of the existing gaze tracking products are wearable, usually with high-quality cameras and infrared devices as active light sources to carry out iris center localization and gaze tracking. Not only is the cost very high, but also the interference to people is great. Therefore, iris center localization in visible light with low-cost devices such as webcams has attracted the interest of more and more researchers. Then, how to accurately locate the iris center in low-quality images with too strong or too weak illuminations, specular reflection, and obscured iris region, is the main issue to deal with in the literature.

## B. IRIS CENTER LOCALIZATION METHODS

In view of the difficulties and challenges of iris center localization, various localization methods are constantly studied, which are mainly divided into three categories: the appearance-based method, the feature-based method and the hybrid method.

1) In the appearance-based method, features do not be extracted explicitly but implicitly, and images are directly mapped to screen coordinates of iris centers. The training model is used to represent the information of eye region indirectly, and the localization is mainly realized by supervised machine learning. Zhang et al. proposed a fast local linear support vector machine (LL-SVM) algorithm to locate the iris center [1]. Bolme et al. proposed the average synthetic precise filter (ASEF) to achieve the location of iris center [2]. Ensemble of random forest is also used in pupil localization [3]. Recently, fully convolutional network is adopted in eye center localization, which has a good localization accuracy for the frontal faces images on BioID and GI4E databases [4], [5]. However, in the appearance-based method, a large number of training samples are often required, and training data under certain conditions are usually adopted. The collected training samples are often dependent on certain persons, so the impact caused by appearance changes cannot be well compensated, and the localization performance is also accidental [6].

2) In the feature-based method, the specific local features of the eye region are mainly identified to locate the iris center. It can be further classified into the method based on shape feature and that based on intensity contrast feature. In the method based on shape feature,

geometric features of eyes are used to establish the model of the eye region. In a simpler model, the iris region is modelled as a circle or oval. In a more complex model, not only is the iris fitted into a circle, but also the upper and lower eyelids are fitted into parabolas. For example, a shape-based method is adopted to fit the pupil into a circle, and then Hough circle transformation is used to achieve pupil position in [7]. Furthermore, two new features of Semi-Circular Edge Shape and Semi-Ellipse Edge Shape are reported in the literature, to make full use of the edge shape features of iris and eyelid [8]. In the method based on shape feature, a high-quality image is often needed. Furthermore, Hough transformation is often used to detect the iris edge in this method, which is very time-consuming. In the method based on gray intensity contrast feature, the gray difference between pupil, iris, and sclera is often employed to locate the iris center. This kind of method is widely used. Daugman proposes a differential integral operator of circular boundary to locate iris center [9]. In the method, the gradient only at the iris boundary is used, while the gray value in the iris center is ignored, so the localization accuracy is low. Valenti et al. propose the localization method based on isophote 8]. They employ the curvature of isophote and curvedness to vote for localization, with a low computational cost, but the accuracy is not too high. Timm et al. propose an iris center location method based on image gradient, with a high localization accuracy in real time [10].

3) In the hybrid method, the feature-based method and the appearance-based method are combined with together to locate the iris center. For example, in [11], gradient-based method and supervised descent method (SDM) are adopted. Firstly, SDM method is used to coarsely estimate the position of the iris center. Then the estimated position is integrated into the gradient-based objective function to get a more accurate iris center location. However, there are also disadvantages of appearance-based method in the hybrid method. Valenti et al. combines the method of isophote with that of machine learning [12]. They adopt a SIFT descriptor and the k-nearest neighbor algorithm to achieve an accurate eye center location, however leading to a high computationally cost. Moreover, Laddi et al. integrate the supervised and unsupervised algorithms, wherein the supervised regression method is used to extract eye regions and the unsupervised gradient method is adopted to localize iris center [13].

By comparing the above methods, in order to meet the requirements of real-time and accurate localization in applications, based on the gradient method proposed by Timm and Barth [10], an iris center localization method using the image inpaint technology, energy map synthesis and post-processing correction is proposed in the paper.

## C. MAIN CONTRIBUTIONS AND INNOVATIONS

In order to improve the accuracy in low-quality images, the image inpaint technology is firstly adopted to inhibit the effect of some specular reflection. Then energy maps based on image gradient and eye ROI midpoint are synthesized to enhance the localization accuracy. In the end, post-processing correction is carried out to eliminate influence of the closed eye and other large derivations to further improve the localization accuracy. The main contributions are summarized as follows.

1) The method of energy map synthesis based on the image gradient and the eye ROI midpoint is proposed to significantly improve the localization accuracy in real time.
2) The image inpaint technology is adopted to inhibit the effect caused by some specular reflection to further improve the localization accuracy.
3) A post-processing correction method is proposed. It makes up for the deficiency of the gradient-based iris center localization method in which the closed eyes cannot be effectively located. Furthermore, the other large deviation of iris center localization caused by the specular reflection, occlusion, black frame glasses, and so on, is corrected to further refine the localization accuracy.

The rest of the paper is organized as follows. First, we introduce the flow chart of the proposed iris center localization system in Section II, and detail the design process of the proposed method mainly including image inpaint, energy map synthesis and post-processing correction in Section III. Then, evaluation method and experimental results are described in Section IV and Section V respectively. Finally, we conclude the work and discuss potential future work in Section VI.

## II. FLOW CHART OF THE PROPOSED IRIS CENTER LOCALIZATION SYSTEM

The flow chart of the proposed iris center localization is shown in Fig. 1. First, the face is detected in the collected image. Then the facial landmark points are detected to get the landmarks around the eye, so that the eye ROI is extracted. Then after the scale normalization transformation of the eye ROI, for one thing, the energy map based on the gradient method is generated after some specular reflection is inhibited by image inpaint; for another, the energy map based on the eye ROI midpoint is generated. Then the two energy maps are synthesized to acquire the final one. The point with the largest gray intensity in the synthetic energy map is found and converted into the position in the original image through the scale inverse transformation, which is the iris center obtained preliminarily. At last, the post-processing correction for the closed eyes and for other large deviations is carried out to achieve an accurate iris center position.
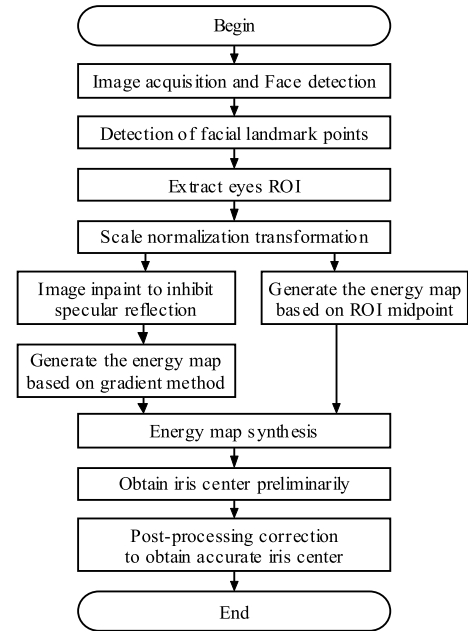


**FIGURE 1.** Flow chart of proposed iris center localization system.

## III. DETAILED DESIGN PROCESS OF LOCALIZATION

### A. FACE DETECTION AND LANDMARKS DETECTION

There are two popular methods for face detection. One is based on the haar-like feature and AdaBoost classifier proposed by Viola and Jones [14], and the other is based on HOG feature and SVM classifier proposed by Dalal and Triggs [15]. In literature [16], the comparison of face detection results is provided between the viola-jones algorithm implemented by OpenCV [17] and the HOG algorithm implemented by Dlib [18]. The results show that for frontal face, the performance of face detection by Dlib exceeds by OpenCV. For the side face, although their detection abilities are limited and their performance is poor, the detection results by Dlib are better than by OpenCV. According to the above comparison, the Dlib algorithm based on HOG is adopted to realize face detection.

The images in the BioID database [19] are tested (see section IV), and two examples of face detection are shown in Fig. 2. Among the 1521 images, the faces in 6 images (146, 485, 638, 949, 1174, 1415) could not be detected, and the left 1515 images were detected successfully. The following experimental results are all based on the 1515 images.
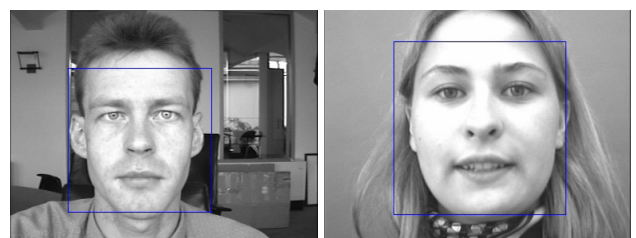


**FIGURE 2.** Two examples of face detection by Dlib.

After face detection, facial landmarks detection is implemented in order to obtain eye ROI. Current facial landmarks detection methods mainly include active shape model (ASM) [20], active appearance model (AAM) [21], restricted local area model (CLM) [22], convolutional neural network (CNN), and ensemble of regression trees (ERT) [23]. ASM, proposed by T.F.Cootes et al, is a deformable model based on statistics, gray and shape separation. AAM is developed based on ASM. Similar to AAM, CLM also combines shape and texture models to produce a model. Unlike AAM, in CLM, a series of local blocks are used. The probability distribution of the landmarks localization is then determined using the mean-shift optimization technique within each local block. The AAM and CNN methods on facial landmarks detection are compared in literature [16]. The results show that the performance of CNN is obviously better than AAM. However, CNN often relies on a large number of input samples and is not appropriate to real-time application. While in the ERT method, the positions of facial landmarks are directly estimated from a sparse subset of pixel intensities, with a better real-time performance and a high prediction quality [23]. In addition, the ERT algorithm is implemented by Dlib library, in which 68 facial landmarks are marked. Therefore, Dlib library is used to realize facial landmarks detection in the paper. Some examples of the detected six landmarks around eyes are shown in Fig. 3. Through the dlib library, we can detect 68 facial landmarks, including six landmarks around one eye. In contrast, in the BioID dataset, only the annotations of 20 facial landmarks are provided, and only two landmarks around one eye are annotated. The localization error of 68 facial landmarks by dlib library on HELEN and LFPW datasets can be found in the original paper [23]. For the localization error of 20 facial landmarks by regression forest on BioID dataset, please refer to Fig.6 in [24].
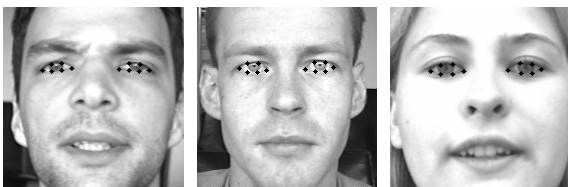


**FIGURE 3.** Examples of detection results of 6 landmarks around human eyes.

### B. EXTRACTION OF EYE ROI AND SCALE NORMALIZATION TRANSFORMATION

According to the detected six landmarks around eyes, the minimum rectangle surrounding them is obtained. Then 1.5 times region of the rectangle is selected as the eye ROI to improve its robustness. The eye ROI hardly includes hair and eyebrows, thus excluding their influence on iris center localization. However, if the subjects wear glasses, especially black-rimmed ones, they will be within the ROI, so their influence on localization will also be taken into account. Some examples of eye ROI are shown in Fig. 4.
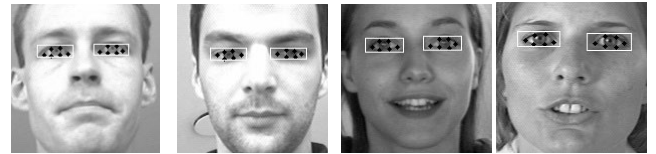


**FIGURE 4.** Examples of eye ROI.

After the extraction of the eye ROI, in order to make the algorithm universal, scale normalization transformation is used. The ratio of transformation is

$$\text{Ratio} = W_{const}/W, \quad (1)$$

where W is the width of the original eye ROI, and $W_{const}$ is the width after normalization, which is set to 50 pixels in the paper. It also seems feasible to normalize the height of eye ROI. However, for the closed-eye images, the height of eye ROI is smaller than its width. Thus, if we normalize the height of eye ROI rather than its width, these images would be larger and the real-time performance would be worse. Moreover, the larger the value of normalization width is, the higher the localization accuracy will be, but the slower the system response will be. However, when its value reaches a certain level, the improvement of localization accuracy will decrease. Therefore, the selection of its value is a trade-off between accuracy and real-time. Under the premise of ensuring a high accuracy, we try to improve the real-time performance. In addition, if the eye ROI is not be set to a fixed size (width or height), which means the eye ROI is not normalized, subsequent image processing and localization will not be carried out on a uniform basis, and the algorithm will not be universal. After the scale normalization transformation of the eye ROI, the iris center will be located according to the following process.

### C. IMAGE INPAINT TO INHIBIT SPECULAR REFLECTION

Due to the specular reflection of light, bright spots will be generated in the iris area, and the intensity in the iris center will increase, leading to the localization deviation of the iris center. In order to inhibit the effect caused by some specular reflection, image inpaint technology is adopted. The image inpaint algorithm of INPAINT_TELEA in OpenCV [25] is used. The process of the image inpaint is shown in Fig. 5.

First, a color image is transformed to gray one after Step1. Afterwards, Gaussian filter is used to eliminate the noise in the image. Then the effects of eyelashes are inhibited by morphological closed operation. Most of the bright spots generated in iris area are similar to ellipse, so ellipse shape structural element is selected. Furthermore, in order not to affect the later localization, a small size of kernel, which is set to 3 here, is used to process the image slightly. After that, the image is segmented by binarization threshold, the results of which directly affect the position and the quality of the image inpaint. The step of binarization threshold segmentation mainly includes the following two processes.
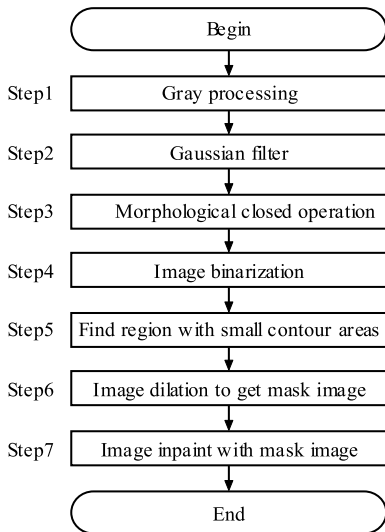
**FIGURE 5.** Flow chart of proposed image inpaint process.

### 1) THRESHOLD SEGMENTATION PROCESS 1

By using the upper limit of the gray threshold, those in the image with higher gray intensity are removed to eliminate the possible eye white area. The binarization processing function of "threshold" in the OpenCV is used, and the binary type of "THRESH _ TOZERO _ INV" is selected to realize the function. The corresponding program statement is: "threshold(eyeGrey, eyeThresh1, thresh1, 255, THRESH_TOZERO_INV);". The relationship between input and output can be expressed as

$$
eyeThresh1 = \begin{cases} 0 & if\,(eyeGrey > thresh1) \\ eyeThresh1 & otherwise. \end{cases} \quad (2)
$$

### 2) THRESHOLD SEGMENTATION PROCESS 2

In the image excluding the possible eye white area, those where the intensity are less than lower limit of gray threshold are set to 0. Others are set to 255 (namely the maximum gray value) to be retained as the possible specular reflection areas. The binary type of "THRESH_BINARY" is chosen to realize the function. The corresponding program statement is: "threshold(eyeThresh1, eyeThresh, thresh2, 255, THRESH_BINARY);". The relationship between input and output can be expressed as

$$
eyeThresh = \begin{cases} 0 & if\,(eyeThresh1 < thresh2) \\ 255 & otherwise. \end{cases} \quad (3)
$$

### 3) DESIGN OF THE THRESH

The upper limit "*thresh1*" and the lower limit "*thresh2*" of the gray threshold are respectively set to

$$thresh1 = m + (M - m) * factor1, \quad (4)$$

$$thresh2 = m + (M - m) * factor2, \quad (5)$$

where *factor1* and *factor2* are coefficients, $m$ and $M$ are the minimum and maximum gray values of the eye ROI

respectively, which have certain adaptability for images from different databases.

The value of *factor1* or *factor2* ranges from 0 to 1, and *factor1* is larger than *factor2*. From (4) and (5), we can see when *factor1* or *factor2* is 0, *thresh1* or *thresh2* equals to $m$ (minimum gray value), and when *factor1* or *factor2* is 1, *thresh1* or *thresh2* equals to $M$ (maximum gray value). The value of *thresh1* or *thresh2* ranges from minimum gray value to maximum one. The schematic diagram of image inpaint area is shown in Fig. 6. The area with gray value greater than *thresh1* is eye white area, and the area with gray value less than *thresh2* is iris area without specular reflection, while the remaining area is the image inpaint area. The value of *factor1* and *factor2* is selected based on experience and on experimental results. The value of *factor1* is set to 0.8, and that of *factor2* is 0.65 in the paper. The formula (4) and (5) ensure that the gray thresholds are in the range of image gray intensities.
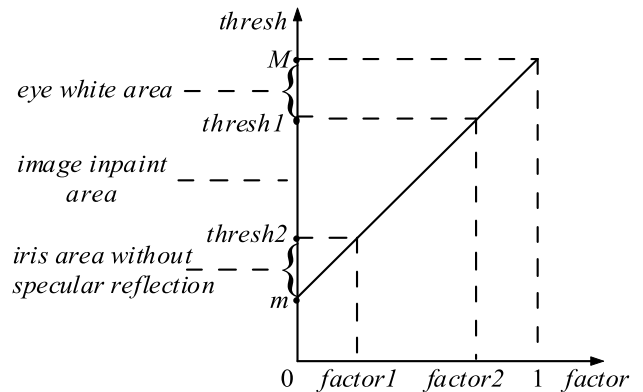


**FIGURE 6.** Schematic diagram of image inpaint area.

After the image binarization, the connected region is extracted to obtain all contours. Then, only those with small area are retained. After that, morphological dilation operation is carried out to acquire a mask image (that is, the specular reflection region). Finally, the image inpaint is conducted to eliminate the influence of some specular reflection by using the mask image. Table 1 shows some examples in the image inpaint process, corresponding to each step in the flowchart of Fig. 5. The leftmost column in Table 1 is the image number. The second column gives the images of the eye ROI before the image inpaint. Moreover, the rightmost column shows the resulting images after the image inpaint. Their localization errors before and after the image inpaint are shown in Table 2. The results show that their localization accuracy has been improved after the image inpaint.

### D. ENERGY MAP AND IRIS CENTER ESTIMATION

After the effect of some specular reflection on the image is inhibited by the image inpaint, the gradient-based energy map will be generated.

**TABLE 1.** Examples of image results in the image inpaint process.

| Image number | Original image | Step1 | Step2 | Step3 | Step4 | Step5 | Step6 | Step7 |
|---|---|---|---|---|---|---|---|---|
| 150 | | | | | | | | |
| 339 | | | | | | | | |
| 489 | | | | | | | | |
| 500 | | | | | | | | |
| 636 | | | | | | | | |
| 818 | | | | | | | | |

**TABLE 2.** Comparison of localization errors before and after inpaint.

| Image number | Localization error before image inpaint | Localization error after image inpaint |
|---|---|---|
| 150 | 0.2525280 | 0.0282335 |
| 339 | 0.0869360 | 0.0217340 |
| 489 | 0.0631320 | 0.0446411 |
| 500 | 0.0644826 | 0.0203912 |
| 636 | 0.0561656 | 0.0311551 |
| 818 | 0.0508183 | 0.0338788 |

**TABLE 3.** Comparison of weight range.

| $I_i$ | $\omega_{ci}=255-I_i$ | $\omega_{ci}=(255-I_i)*k+b$ |
|---|---|---|
| 0 | 255 | 24 |
| 255 | 0 | 20 |

### 1) ENERGY MAP BASED ON IMPROVED GRADIENT

The formula of iris center localization based on the gradient is given by [10]

$$
\begin{cases}
\hat{c} = \arg\max_c \left\{ \dfrac{1}{n} \sum_{i=1}^{n} \omega_{ci}(d_i^T g_i)^2 \right\} \\
d_i = \dfrac{p_i - c}{\|p_i - c\|_2}, \quad \forall i : \|g_i\|_2 = 1,
\end{cases} \tag{6}
$$

where, $\hat{c}$ is the estimated iris center, $c$ is the possible iris center, $g_i$ is the gray gradient vector at a point $p_i$ of the iris edge, $d_i$ is the displacement vector from $c$ to $p_i$, $n$ is the total number of pixels in the image. The weight value $\omega_{ci}$ of $c$ is given by

$$
\omega_{ci} = 255 - I_i, \tag{7}
$$

where $I_i$ is gray value at $c$. Under normal circumstances, the image near the iris center is darker, with a smaller gray intensity value, and the edge of the iris is just the opposite, so the gray value can be inverted as the weight just as (7). However, due to the influence of some specular reflection, bright spots are often generated in the iris area, leading to the increase of gray intensity near the iris center. Though the effect of some specular reflection is inhibited by the image inpaint, we cannot completely eliminate their effects. To this end, the weight value can be redesigned as

$$
\omega_{ci} = (255 - I_i)*k + b, \tag{8}
$$

where $k$ is a coefficient less than 1, and $b$ is a small constant. The value of $k$ is set to 0.015, and $b$ is 20, which are determined by experiment results. The Comparison of weight range is shown in Table 3. It can be seen that the range of weights varies from 0-255 to 20-24. The redesigned weight weakens the effect of image gray value, in order to suppress the effect of some specular reflection.

According to (8) and the curly brace portion in the first line of (6), the gray value of each point in the eye ROI can be acquired, whereby the energy map based on the gradient is generated. The images in the BioID database are tested (see section IV). Some examples of the energy map based on the above gradient method are shown in the second column of Table 4.

**TABLE 4.** Synthesis process of energy map.

| Image number | Energy map based on gradient method | Energy map based on ROI midpoint | Synthesized energy map |
|---|---|---|---|
| 53 | | | |
| 168 | | | |
| 346 | | | |
| 592 | | | |
| 658 | | | |
| 727 | | | |
| 863 | | | |
| 1139 | | | |
| 1182 | | | |
| 1295 | | | |

### 2) PROPOSED ENERGY MAP BASED ON THE EYE ROI MIDPOINT

Because the iris center is close to the eye ROI midpoint, the energy map based on the eye ROI midpoint is generated to correct that based on the gradient. This kind of energy map is composed of a white circle with blurred edges. The eye ROI midpoint is the center of the circle, and its gray value is maximum with 255; the gray value decreases gradually within a certain radius, and the gray value at the edge of the circle is minimum with 0. Therefore, the energy map can be represented by a two-dimensional Gaussian function. However, the function value ranges from 0 to 1, while the gray

value of an image ranges from 0 to 255. Thus, the energy map based on the eye ROI midpoint can be given by
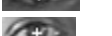
$$E_m(x, y) = 255 * e^{-\frac{(x-x_m)^2+(y-y_m)^2}{\sigma^2}}, \qquad (9)$$

where $(x_m, y_m)$ is the coordinate of the eye ROI midpoint, $\sigma$ is the radius of the circle in the energy map, which is set to 6 here. If the value of $\sigma$ is too small, the effect of the eye ROI midpoint will be too large, and the localization error will be increased. Moreover, if $\sigma$ is too large, the effect of the eye ROI midpoint will be weakened, its correctness effect on the gradient-based energy map will become worse. The value of $\sigma$ is determined according to the experiment results in the paper. Some examples of energy map base on the eye ROI midpoint are shown in the third column of Table 4.

### 3) PROPOSED ENERGY MAP SYNTHESIS AND IRIS CENTER ESTIMATION

After the energy maps based on the gradient and the eye ROI midpoint are generated, they will be synthesized to obtain the final one. The image fusion method can be used for synthesis, and the premise of which is that the two energy maps have the same size. Since both energy maps are all based on the normalized size of the eye ROI, the fusion condition is satisfied. They can be synthesized according to a certain weight ratio, in which the energy map based on the gradient occupies a higher proportion, while that based on the eye ROI midpoint takes up a smaller weight. For the images listed in the leftmost of Table 4, the synthesized energy maps are shown in its rightmost column. In the synthesized energy map, the point with the maximum gray intensity is the estimated iris center. The algorithm is shown in Algorithm 1 of Appendix. For the images in Table 4, the localization results of the iris center before and after the energy map synthesis are shown in Table 5, and their localization errors are shown in Table 6. The results show that their localization accuracy has been greatly improved.

**TABLE 5.** Comparison of localization results before and after energy map synthesis.

| Image number | Localization results before energy map synthesis | Localization results after energy map synthesis |
|---|---|---|
| 53 | | |
| 168 | | |
| 346 | | |
| 592 | | |
| 658 | | |
| 727 | | |
| 863 | | |
| 1139 | | |
| 1182 | | |
| 1295 | | |

**TABLE 6.** Comparison of localization errors before and after energy map synthesis.

| Image number | Localization errors before energy map synthesis | Localization errors after energy map synthesis |
|---|---|---|
| 53 | 0.0566243 | 0.0111049 |
| 168 | 0.2431160 | 0.0446856 |
| 346 | 0.3458760 | 0.0363576 |
| 592 | 0.0526235 | 0.0392232 |
| 658 | 0.1745810 | 0.0354887 |
| 727 | 0.0686803 | 0.0471143 |
| 863 | 0.0928339 | 0.0344776 |
| 1139 | 0.0522352 | 0.0184679 |
| 1182 | 0.1791170 | 0.0222167 |
| 1295 | 0.1649290 | 0.0184396 |

### E. PROPOSED POST-PROCESSING CORRECTION ALGORITHM

When the eyes are closed, it is difficult to obtain accurate iris center locations by using gradient-based method, so the positions of the iris center will be corrected. Furthermore, as we know, the iris center is close to the eye ROI midpoint and far from the eye ROI boundary, and the interpupillary distance is about two times of the eye width. Therefore, when the estimated iris center is too close to the boundary of the eye ROI, or interpupillary distance is too large or too small, it is obvious that the localization error is large, and the position of the iris center will also be corrected. The proposed post-processing correction algorithm is shown in Algorithm 2 of Appendix.

### 1) DETECTION AND CORRECTION OF THE CLOSED EYE

According to prior knowledge, when the eye are closed, its ROI is relatively flat, i.e., the ratio of width to height is relatively large. Accordingly, the state of the closed eyes can be judged. In the closed eye state, the positions of iris center of both eyes are corrected with the midpoints of their ROIs.

### 2) THE DISTANCE FROM THE ROI BOUNDARY IS TOO SMALL

When the estimated iris center is too close to the inner boundary (*edgeInner*) of the ROI, its outer boundary (edge-Outer), its upper boundary (*edgeUpper*), or its lower boundary (*edgeLower*), it shows that the localization error of the iris center is large, so its location will be corrected. The limit of distance between the iris center and the boundary of the ROI is *dmin*, which is set according to the subsequent formula (10).

### 3) INTERPUPILLARY DISTANCE IS TOO LARGE OR TOO SMALL

When the interpupillary distance (*DistancePupils*) is too small, the position of the iris center will be corrected. If the iris center of the left eye is close to its inner boundary (*edgeInner*), which indicates that the localization error of the left eye is relatively large and its iris center will be corrected, so its correction flag (*flagLeft*) is set to 1. On the contrary, if the iris center of the right eye is close to its inner boundary,

its correction flag (*flagRight*) is set to 1. When the interpupillary distance is too large, the position of the iris center will also be corrected. If the iris center of the left eye is close to its outer boundary (*edgeOuter*), its iris center will be corrected. On the contrary, if the iris center of the right eye is close to its outer boundary, its iris center will be corrected. When the *flagLeft* is 1, the midpoint ($xLeft_c$, $yLeft_c$) of the left eye ROI will be adopted as its iris center. Similarly, when the *flagRight* is 1, the midpoint ($xRight_c$, $yRight_c$) of the right eye ROI will be used as its iris center.

The parameters of the correction are set according to prior knowledge. As stated above, the ROI is selected as 1.5 times size of minimum rectangle region surrounding the landmark points in the eye region. If the width of the ROI is $W$, then the width of the eye is $2W/3$, and the distance between the midpoint of the eye and its corner is $W/3$, and the distance between the corner of the eye and the edge of the ROI is $W/6$, as shown in Fig. 7. When the distance between the iris center and the edge of the ROI is less than *dmin*, the position of the iris center will be corrected. From Fig. 7, we have

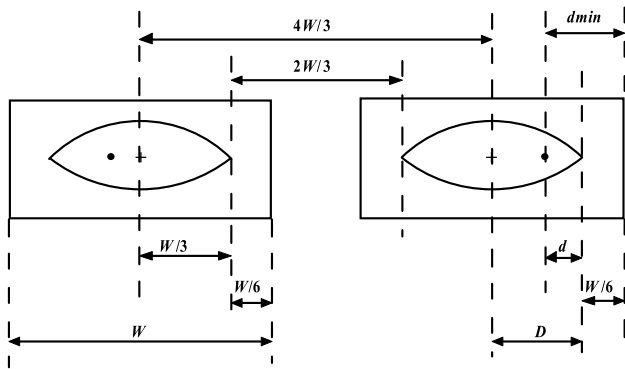$$dmin = d + W/6, \qquad (10)$$



**FIGURE 7.** Parameter diagram of proposed post-processing correction.

where $d$ is the distance between the iris center and the nearest corner of the eye. According to the experience, we know normally

$$d \geq D/10, \qquad (11)$$

where $D = W/3$, so we can derive that $dmin = 0.2W$. Likewise, we can set the threshold of distance between the iris center and its upper or lower edge. Furthermore, as we know, the interpupillary distance is about two times size of the eye width, so the interpupillary distance is about $4W/3$. When the interpupillary distance is smaller than a low threshold (namely *din*) or larger than a high threshold (namely *dout*), the position of the iris center will be corrected. In order to enhance the robustness, a certain margin is left, i.e. *din* is set to $1.2W$, while *dout* is set to $1.5W$ here. When the interpupillary distance is too large or too small, it is obvious that the position of the iris center is incorrect. In this case, the minimum distance (namely *dmid*) between the iris center and the edge of the ROI can be set to a bigger value, such as $0.26W$ here, to acquire an accurate location.

Some examples of localization results before and after the post-processing correction are shown in Table 7. The iris centers are marked with cross lines. It can be seen that in the closed-eye state or when the large deviations appear, caused by specular reflection, black-rimmed glasses, and so on, the positions of the iris centers are all corrected. The localization errors before and after the post-processing correction are compared in Table 8, which shows that the localization accuracy after correction has been further improved.

**TABLE 7.** Comparison of localization results before and after post-processing correction.

| Image number | Localization result before correction | Localization result after correction |
|---|---|---|
| 158 | | |
| 167 | | |
| 327 | | |
| 594 | | |
| 744 | | |
| 1016 | | |

**TABLE 8.** Comparison of localization errors before and after post-processing correction.

| Image number | localization error before correction | localization error after correction |
|---|---|---|
| 158 | 0.228123 | 0.0408078 |
| 167 | 0.259533 | 0.0446411 |
| 327 | 0.122359 | 0.0454428 |
| 594 | 0.290708 | 0.0372213 |
| 744 | 0.144787 | 0.0476056 |
| 1016 | 0.0725238 | 0.0362619 |

## IV. EVALUATION OF IRIS CENTER LOCALIZATION PERFORMANCE

The quality of the localization results needs to be verified on face databases, as well as good evaluation indexes and evaluation methods needed. The algorithm can be further optimized according to the feedback location results. The dataset selection and performance evaluation are as follows.

### A. DATASET SELECTION

The public face database can be used to measure the algorithm, and the more the database approaches to the real world, the more accurately it will measure. Thus the BioID database [19] can be the most ideal choice. The database contains 23 different subjects and 1,521 gray images with a resolution of 384*286. Moreover, it also contains some closed eye and semi-closed eye images, and even those in which eye area is blocked by the strong light of the glasses, so it is considered publically the most challenging database. Furthermore, it also provides true iris center location of both eyes, which provides the opportunity for the localization.

### B. EVALUATION INDEXES AND METHOD

#### 1) EVALUATION INDEXES

Whether the localization results are precise requires certain evaluation indexes. The maximum normalized localization

error is often used as an evaluation index for iris center localization [26], which is given by

$$e = \frac{\max(e_l - e_r)}{d_{lr}} = \frac{\max(\|\hat{c}_l - c_l\|_2, \|\hat{c}_r - c_r\|_2)}{\|c_l - c_r\|_2}, \quad (12)$$

where $\hat{c}_l$ and $\hat{c}_r$ are the estimated iris center of left eye and right eye respectively; $c_l$ and $c_r$ are their actual iris center respectively; $e_l$ and $e_r$ represent the Euclidean distance between the estimated position and the ground truth, namely their absolute localization error; $d_{lr}$ represents the actual Euclidean distance between the iris centers of both eyes. The formula (12) represents the ratio of the maximum value of absolute localization errors of both eyes to the distance between them. It can be seen that the localization error is independent of image size and face size, which is normalized. In addition, minimum normalization error and average one are also used as the evaluation indexes.

Moreover, e ≤ 0.05, e ≤ 0.10 and e ≤ 0.25 respectively indicate that the localization error is in the range of pupil, iris and eye center to eye corner.

### 2) EVALUATION METHODS
In order to further improve the algorithm, good evaluation methods are desirable.

#### a: PROCESSING OF INPUT IMAGE
If each image in the BioID database is taken as the input separately, the workload is very heavy. Therefore, a real-time video acquisition method is adopted. All 1521 images in the database are taken as video frames for sequential acquisition, followed by identification and localization. Then the localization result of the iris center for each image is stored as a row of data, and the image number is saved in front of the localization result. The localization results for 1515 detected images are all saved in one file.

#### b: EXTRACTION OF ACTUAL EYE POSITION
In order to locate all the images in the database automatically, the true values of the iris centers for all 1521 images need to be extracted automatically. However, for each image, its actual eye position is stored in an individual file. Therefore, it is desirable to extract the true values of the eye positions of all images successively and save them in one file, which is mainly realized by deserialization and serialization.

#### c: EVALUATION OF LOCALIZATION PERFORMANCE
The data is imported from the estimated iris center localization file and the actual eye position file. Then the localization errors of the iris centers are computed, and the results are assessed by using the above evaluation indexes. At the same time, the proportion of the corresponding image in each range of the errors is counted to obtain the localization accuracy. Finally, the localization error and localization accuracy are saved in the output file.

### 3) RESULTS ANALYSIS
The above evaluation method is adopted to get the maximum normalization localization error and localization accuracy of 1515 images detected in the BioID database. The statistical results of localization accuracy are shown in Table 9.

**TABLE 9.** Localization accuracy of the iris center in each stage.

| Method | e≤0.05 | e≤0.1 | e≤0.25 |
|---|---|---|---|
| Basic gradient method | 80.3% | 91.0% | 96.8% |
| After image inpaint | 81.7% | 92.0% | 97.6% |
| After energy map synthesis | 87.6% | 95.3% | 98.6% |
| After post-processing correction | 89.6% | 98.7% | 99.9% |

The first row of data is the localization accuracy of the iris center by using the basic gradient method; the second row is that after the image inpaint; the third one is that after the energy map synthesis; and the last one is that after the post-processing correction. By comparing the first two rows of data in Table 9, we can see that after the image inpaint, the localization accuracy is increased by 1.4% when the localization error is in the range of 0.05; that is raised by 1.0% when the error is within 0.1. By the contrast of the data in the second and third row, it can be seen that after the energy map synthesis, when the error is within 0.05, the localization accuracy is improved by 5.9%; when the error is within 0.1, that is increased by 3.3%. It is obvious that the energy map synthesis significantly improves the localization accuracy. From the last two rows of data, we can see that after the post-processing correction, when the error is within 0.05, the localization accuracy is improved by 2.0%; when the error is within 0.1, that is increased by 3.4%. After the image inpaint, the energy map synthesis and the post-processing correction, the localization accuracy is improved by 9.3% when the error is in the range of 0.05; increased by 7.7% when the error is within 0.1; raised by 3.1% when the error is within 0.25. The localization accuracy corresponding to the maximum, average and minimum normalized errors is shown in Table 10, and their curves are shown in Fig. 8. By comparing the first two rows of data in Table 10, we can see that the localization accuracy in average normalized error is 5.8% higher than that in maximum normalized error when the error is within 0.05, 0.8% higher within 0.1. By comparing the first row and the last row of data in Table 10, it can be seen that the localization accuracy in the minimum normalized error is 8.8% higher than that in the maximum normalized error when the error is within 0.05, 1.2% higher within 0.1. It is clear that in the

**TABLE 10.** Localization accuracy corresponding to maximum, average and minimum normalized errors.

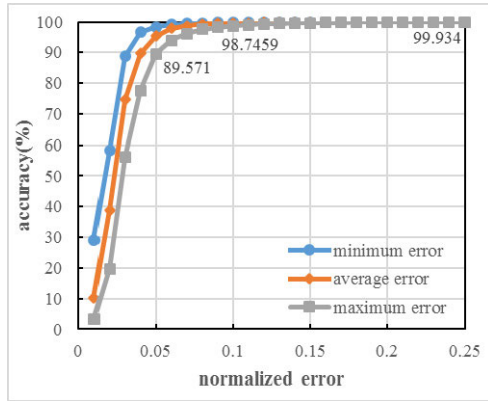| Method of measuring localization accuracy | e≤0.05 | e≤0.1 | e≤0.25 |
|---|---|---|---|
| Maximum normalized error | 89.6% | 98.7% | 99.9% |
| Average normalized error | 95.4% | 99.5% | 99.9% |
| Minimum normalized error | 98.4% | 99.9% | 99.9% |

**FIGURE 8.** Localization accuracy curve.

case of average and minimum normalization errors, a higher localization accuracy and a better performance are achieved.

The localization accuracy of the proposed method on the BioID database is compared with other advanced methods, as shown in Table 11.

**TABLE 11.** Comparison of localization accuracy on BioID database.

| Method | e≤0.05 | e≤0.1 | e≤0.25 |
|---|---|---|---|
| Timm2011[10] | 82.5% | 93.4% | 98.0% |
| Valenti2008[27] | 84.1% | 90.9% | 98.5% |
| Valenti2012[12] | 86.1% | 91.7% | 97.9% |
| Markuš2014[3] | 89.9% | 97.1% | 99.7% |
| Leo2014[28] | 80.7% | 87.3% | 93.9% |
| Zhang2016[29] | 85.7% | 93.7% | 99.2% |
| Laddi2017[30] | 81.4% | 92.2% | 97.5% |
| Cai2015[31] | 84.1% | 95.6% | 99.8% |
| Cai2017[32] | 86.8% | 96.6% | 99.9% |
| Xia2018[11] | 87.1% | 98.7% | 99.9% |
| Xia2019[4] | **94.4%** | **99.9%** | **100%** |
| Choi2019[5] | 93.3% | -- | -- |
| Laddi2019[13] | 81.4% | 92.2% | 97.5% |
| Ahmed 2019[8] | 87.0% | 95.2% | 99.0% |
| Proposed method | <u>89.6%</u> | <u>98.7%</u> | <u>99.9%</u> |

In different error ranges, the results with the highest localization accuracy are marked in bold. The localization accuracy of the proposed method is listed in the last row, marked with underline. The localization accuracy of the proposed method is higher than any other methods listed in Table 11 except [3]–[5]. Though they achieve higher accuracies than our method on the BioID database, they have their limitations and disadvantages. The comparison between our approach and their methods is summarized as follows. Firstly, because their methods are supervised learning approaches, they are very dependent on training samples and prone to overfitting, with a poor generalization ability. Their prediction accuracies beyond the range of training data will degrade greatly. In contrast, the proposed method is an unsupervised approach without the need of training samples, whose generalization ability and adaptability are much better. Secondly, their methods are tested only in the frontal face databases such as BioID or GI4E databases. They are hard to adapt to

variable pose images, with poor robustness on pose variation. In contrast, the proposed method has a better robustness to different pose images, which has been verified on the Talking Face Video database and the MUCT face database (see section V). Thirdly, their methods normally need the support of advanced computer hardware such as GPU and large RAM memory. Furthermore, the time cost for training will be very expensive with lower levelled computer hardware. The consumed time for testing will be prolonged. In contrast, the proposed method only requires the support of hardware such as CPU and small RAM memory, without the need of GPU. We have achieved a good real time performance with lower levelled computer hardware (see the subsequent real time performance). Moreover, the real time performance will be improved significantly in the support of advanced computer hardware. From Table 11, it can be seen that the localization accuracy of the proposed method has outperformed the state-of-the-art unsupervised approaches on the BioID database.

The localization accuracy of the iris center estimated by basic gradient method (see the first row of data in Table 9) is not as high as that in [10] (see the first row of data in Table 11), which indicates that the gradient method used in the paper is not very perfect or some parameter values are not the most optimal. However, after the subsequent image inpaint, the energy map synthesis and the post-processing correction, the localization accuracy is significantly improved, greatly higher than that in [10], which also shows that the robustness of the proposed method is strong.

In addition, in terms of real time performance, the proposed method is compared with the other approaches as shown in Table 12. As the proposed method is based on the gradient method, it has not the disadvantage of poor real-time performance of regression or machine learning algorithm. On the PC with Intel Core i5-3.2GHz CPU and 4G RAM, the average consumed time of localization algorithm for the 1515 detected images in the BioID database is about 50ms, i.e. with the speed of 20 frames per second, which shows that the algorithm has a good real-time performance.

**TABLE 12.** Comparison of real time performance.

| Method | Consumed time(ms) | Speed(frame/s) |
|---|---|---|
| Leo2014[28] | 333 | 3 |
| Araujo2014[33] | 83 | 12 |
| Eddine2018[34] | 161 | 6 |
| Poulopoulos2019[35] | 50-62.5 | 16-20 |
| Proposed method | 50 | 20 |

## V. ROBUSTNESS TO POSE AND ILLUMINATION

Since BioID database mainly consists of frontal faces, it is important to evaluate the robustness of the algorithm to various poses and variable illuminations. Thus, in order to verify the robustness of the algorithm, the method is tested on the other two databases of Talking Face Video database [36] and the MUCT face database [37]. Talking Face Video database consists of 5000 images of a person engaged in conversation,

with different poses and various expressions. The MUCT face database consists of 3755 images with multiple poses in variable illuminations. Among the 5000 Talking Face images, all faces were detected successfully. While among the 3755 MUCT face images, 3752 faces were detected successfully. The following experimental results are all based on the above detected images. Some examples of the detected face regions, eye ROIs and corresponding localization results on the Talking Face Video database are shown in Fig. 9. Their iris center positions are represented with white cross line in eye ROI images. Fig. 10 displays some examples of iris center localization results on the MUCT face database.
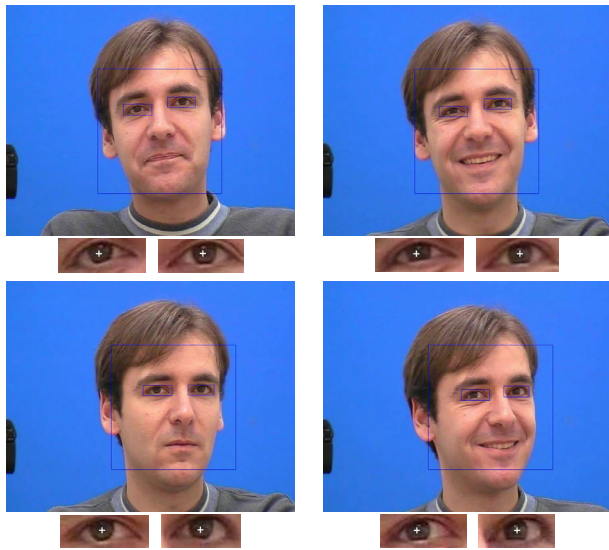


**FIGURE 9.** Examples of iris center localization results on Talking Face Video database.



**FIGURE 10.** Examples of iris center localization results on the MUCT face database.

The localization accuracy of the iris center at each stage on Talking Face Video database is shown in Table 13.

**TABLE 13.** Localization accuracy of the iris center at each stage on talking face video database.

| Method | e≤0.05 | e≤0.1 | e≤0.25 |
|---|---|---|---|
| Gradient method | 83.7% | 96.9% | 99.8% |
| After energy map synthesis | 93.6% | 98.4% | 99.8% |
| After post-processing correction | 94.3% | 99.8% | 99.98% |

The images in this database are of high quality, so there is no need for image inpaint. Compared with the gradient method, when the normalized error is within 0.05, after energy map synthesis, the localization accuracy of the proposed method is improved by 9.9%; after post-processing correction, increased by 0.7%; totally improved by 10.6%. When the normalized error is within 0.1, the localization accuracy is totally improved by 2.9%. Table 14 displays the localization accuracy on the MUCT face database. From the table, we can see that for the gradient method, the localization accuracy is very low, because of the huge effects of the dark eyelashes, black-rimmed glasses or dark skin in many images. While, when the normalized error is within 0.05, after image inpaint, the localization accuracy is improved by 1.8%; after energy map synthesis, improved by 13.8%; after post-processing correction, increased by 2.1%; improved by 17.7% totally compared with the gradient method. Compared with the gradient method, the localization accuracy of the proposed method is improved by 8.4% when the normalized error is within 0.1, increased by 1.5% within 0.25.

**TABLE 14.** Localization accuracy of the iris center at each stage on muct face database.

| Method | e≤0.05 | e≤0.1 | e≤0.25 |
|---|---|---|---|
| Gradient method | 66.9% | 88.8% | 98.4% |
| After image inpaint | 68.7% | 89.5% | 98.5% |
| After energy map synthesis | 82.5% | 94.9% | 99.1% |
| After post-processing correction | 84.6% | 97.2% | 99.9% |

Comparisons of localization accuracy with the other advanced methods on the two databases are shown in Table 15 and Table 16 respectively. It can be seen that the localization performance outperforms the other methods listed in the tables. Furthermore, the results in the two databases verify that the proposed algorithm is robust to the pose and illumination.

**TABLE 15.** Comparison of localization accuracy on talking face video database.

| Method | e≤0.05 | e≤0.1 | e≤0.25 |
|---|---|---|---|
| Pang2015[38] | -- | 96.2% | -- |
| Ahmed 2019[8] | -- | 98.8% | -- |
| Laddi2019[13] | 90.2% | 98.7% | 99.5% |
| Proposed method | 94.3% | 99.8% | 99.98% |

**TABLE 16.** Comparison of localization accuracy on muct face database.

| Method | e≤0.05 | e≤0.1 | e≤0.25 |
|---|---|---|---|
| Laddi2019[13] | 78.6% | 88.3% | 92.2% |
| Proposed method | 84.6% | 97.2% | 99.9% |

## VI. CONCLUSION AND PROSPECT

In order to accurately locate the iris center in low-quality images and meet the need of efficiency in practical applications, a novel method of iris center localization is proposed in this paper using energy map synthesis based on image gradient, image inpaint technology, and post-processing correction. Moreover, the algorithm is verified on the BioID database. The image inpaint technology is firstly adopted to inhibit the effect of some specular reflection to improve the localization accuracy of the iris center. Then the energy maps based on image gradient and the eye ROI midpoint are synthesized to improve localization accuracy significantly. Especially localization accuracy is enhanced by 5.9% in the small range of maximum normalization localization error ($e \leq 0.05$). After that, a post-processing correction method is proposed and used, in order to further refine the localization accuracy. Particularly the localization accuracy is improved by 3.4% in the range of larger localization error ($e \leq 0.1$). Overall, after the image inpaint, the energy map synthesis and the post-processing correction, the localization accuracy is 89.6% (improved by 9.3%) when the localization error is within 0.05, that reaches 98.7% (improved by 7.7%) when the localization error is within 0.1. Compared with the other advanced methods on the BioID database, the proposed method has outperformed the state-of-the-art unsupervised methods, which has also confirmed the validity of this method. Furthermore, the algorithm has high real-time performance with the speed of 20 frames per second, which has laid a foundation for its further applications in gaze tracking and other aspects.

Moreover, the algorithm is tested on the other two databases of Talking Face Video database and the MUCT face database, which has verified the robustness to various poses and variable illuminations. Comparing the results of image inpaint, energy map synthesis and post-processing correction in the proposed method, the localization accuracy has been improved significantly by energy map synthesis, while image inpaint and post-processing correction are helpful to refine localization accuracy in further. However, the proposed method has also limitations, because face detection and landmark detection before iris center localization are based on dlib, though the consumed time for localization algorithm is short, the whole running time is prolonged. If dlib is used, the consumed time is short in the Ubuntu operating system, while it will take more time in Windows systems. Thus, the proposed localization algorithm is more applicable to run in Ubuntu systems. The other alternative is to run the proposed method based on other fast face detection and landmarks detection algorithms, which we will explore and research in the future.

In addition, the localization accuracy of the iris center based on basic gradient method is compared with that obtained by Timm, there is still a room for improvement, so in the future we will upgrade the program to further improve localization accuracy.

In the future, based on the iris center localization, we will research on gaze tracking and iris biometrics of astronauts by virtue of astronaut assistant robot. Mobile device is installed on the astronaut assistant robot to capture the images of astronauts by the camera embedded in the mobile device. For one thing, gaze tracking can be employed to detect the health state of astronauts including physical and psychological health. For another, iris biometrics can be applied to identify astronauts in order to make them have access to do some special jobs. Furthermore, the mobile device is widely used by almost everyone. In the near future, it is likely that

---

**Algorithm 1** Energy Map Synthesis Algorithm

**Input**: **ImageOfEyeROI MidpointOfEyeROI** $(x_m, y_m)$
**Initialize**: **parameters** k,m, $\sigma$, W
**Compute**: **DisplacementVector** $d_i$
   **GradientVector** $g_i$

$$d_i = (x_i - x_c, y_i - y_c)^T$$
$$\|d_i\|_2 = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}$$
$$d_i = \frac{d_i}{\|d_i\|_2}$$
$$g_i = \left( \frac{\partial I_i}{\partial x_i}, \frac{\partial I_i}{\partial y_i} \right)^T$$
$$\|g_i\|_2 = \sqrt{\left( \frac{\partial I_i}{\partial x_i} \right)^2 + \left( \frac{\partial I_i}{\partial y_i} \right)^2}$$
$$g_i = \frac{g_i}{\|g_i\|_2}$$

**Compute**: **WeightValue** $\omega_{ci}$

$$\omega_{ci} = (255 - I_i)^* k + m$$

**Please refer to formula(8)**

**Compute**: **EnergyMapOfGradient** $E_{gi}$
   **EnergyMapOfEyeROImidpoint** $E_{mi}$
   **SyntheticEnergyMap** $E_i$

$$E_{gi} = \frac{1}{n} \sum_{i=1}^{n} \left[ \omega_{ci} \left( d_i^T g_i \right)^2 \right]$$

**Please refer to formula(6)**

$$E_{mi} = 255 \cdot e^{\frac{(x_i - x_m)^2 + (y_i - y_m)^2}{\sigma}}$$

**Please refer to formula(9)**

$$E_i = W E_{gi} + (1 - W) E_{mi}$$

**Compute**: **IrisCenter** $(x_c, y_c)$

$$(x_c, y_c) = \max_{i=1}^{n} E_i$$

**Output**: **IrisCenter** $(x_c, y_c)$

---

it will be widely applied to capture iris images of people and identify them [39]. The MICHE iris database [40], in which the images are collected by mobile devices, will be used to verify and test the related algorithms. To sum up, the iris center localization lays a good foundation for the future gaze tracking and iris biometrics.

## APPENDIX
See Algorithms 1 and 2.

---

**Algorithm 2** Postprocessing Correction

---

**Input**: *originCenterLeft, originCenterRight,    ROI*
**Compute**: *edgeInnerLeft, edgeOuterLeft, edgeUpperLeft*
      *edgeLowerLeft*
      *edgeInnerRight, edgeOuterRight,*
      *edgeUpperRig ht, edgeLowerRight*
      *dis tan cePupils*
**if**  *widthROI/heightROI > ratio*  **then**
    *flagLeft = 1*
    *flagRight = 1*
**end if**
**if**
    *(edgeInnerLeft < dmin) or (edgeOuterLeft < dmin)*
    *or (edgeUpperLeft < dmin) or (edgeLowerLeft < dmin2)*
    *flagLeft = 1*  **then**
**end if**
**if**
  *(edgeInnerRight < dmin) or (edgeOuterRight < dmin)*
  *or (edgeUpperRight < dmin2) or (edgeLowerRight < dmin2)*
  *flagRight = 1*  **then**
**end if**
**if**  *distancePupils < din*  **then**
    **if**  *edgeInnerLeft < dmid*  **then**
      *flagLeft = 1*
    **end if**
    **if**  *edgeInnerRight < dmid*  **then**
      *flagRight = 1*
    **end if**
**end if**
**if**  *distancePupils > dout*  **then**
    **if**  *edgeOuterLeft < dmid*  **then**
      *flagLeft = 1*
    **end if**
    **if**  *edgeOuterRight < dmid*  **then**
      *flagRight = 1*
    **end if**
**end if**
**if**  *flagLeft = 1*  **then**
    *CenterLeft = (xLeft$_c$, yLeft$_c$)*
**end if**
**if**  *flagRight = 1*  **then**
    *CenterRight = (xRight$_c$, yRight$_c$)*
**end if**
**Output**:    CenterLeft, CenterRight

---

## REFERENCES

[1] C. Zhang, X. Sun, J. Hu, and W. Deng, "Precise eye localization by fast local linear SVM," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2014.

[2] D. S. Bolme, B. A. Draper, and J. R. Beveridge, "Average of synthetic exact filters," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 2105–2112.

[3] N. Markuš, M. Frljak, I. S. Pandžić, J. Ahlberg, and R. Forchheimer, "Eye pupil localization with an ensemble of randomized trees," *Pattern Recognit.*, vol. 47, no. 2, pp. 578–587, Feb. 2014.

[4] Y. Xia, H. Yu, and F.-Y. Wang, "Accurate and robust eye center localization via fully convolutional networks," *IEEE/CAA J. Autom. Sinica*, vol. 6, no. 5, pp. 1127–1138, Sep. 2019.

[5] J. H. Choi, K. I. Lee, Y. C. Kim, and B. Cheol Song, "Accurate eye pupil localization using heterogeneous CNN models," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 2179–2183.

[6] S. Cristina and K. P. Camilleri, "Unobtrusive and pervasive video-based eye-gaze tracking," *Image Vis. Comput.*, vol. 74, pp. 21–40, Jun. 2018.

[7] M. T. Setiawan, S. Wibirama, and N. A. Setiawan, "Robust pupil localization algorithm based on circular Hough transform for extreme pupil occlusion," in *Proc. 4th Int. Conf. Sci. Technol. (ICST)*, Aug. 2018, pp. 1–5.

[8] M. Ahmed and R. H. Laskar, "Eye center localization in a facial image based on geometric shapes of iris and eyelid under natural variability," *Image Vis. Comput.*, vol. 88, pp. 52–66, Aug. 2019.

[9] J. G. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 11, pp. 1148–1161, Nov. 1993.

[10] F. Timm and E. Barth, "Accurate eye centre localisation by means of gradients," in *Proc. Int. Conf. Comput. Vis. Theory Appl. (VISAPP)*, 2011, pp. 125–130.

[11] Y. Xia, J. Lou, J. Dong, G. Li, and H. Yu, "SDM-based means of gradient for eye center localization," in *Proc. IEEE 16th Int. Conf. Dependable, Autonomic Secure Comput., 16th Int. Conf. Pervasive Intell. Comput., 4th Int. Conf Big Data Intell. Comput. Cyber Sci. Technol. Congr. (DASC/PiCom/DataCom/CyberSciTech)*, Aug. 2018, pp. 862–867.

[12] R. Valenti and T. Gevers, "Accurate eye center location through invariant isocentric patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1785–1798, Sep. 2012.

[13] A. Laddi and N. R. Prakash, "Eye gaze tracking based directional control interface for interactive applications," *Multimedia Tools Appl.*, vol. 78, no. 22, pp. 31215–31230, Nov. 2019.

[14] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, May 2004.

[15] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jul. 2005, pp. 886–893.

[16] B. Johnston and P. de Chazal, "A review of image-based automatic facial landmark identification techniques," *EURASIP J. Image Video Process.*, vol. 2018, no. 1, p. 86, 2018.

[17] O. Team. (2018). *OpenCV*. [Online]. Available: https://opencv.org/

[18] D. E. King, "Dlib-ml: A machine learning toolkit," *J. Mach. Learn. Res.*, vol. 10, pp. 1755–1758, Jul. 2009.

[19] (2019). *The BioID Face Database*. [Online]. Available: http://www.bioid.com

[20] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models-their training and application," *Comput. Vis. Image Understand.*, vol. 61, no. 1, pp. 38–59, Jan. 1995.

[21] G. Edwards, C. Taylor, and T. Cootes, "Interpreting face images using active appearance models," in *Proc. 3rd IEEE Int. Conf. Autom. Face Gesture Recognit.*, Nov. 2002, pp. 300–305.

[22] D. C. A. T. Cootes, "Feature detection and tracking with constrained local models," in *Proc. Brit. Mach. Vis. Conf.*, Edinburgh, U.K., 2006.

[23] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1867–1874.

[24] H. Yang and I. Patras, "Privileged information-based conditional structured output regression forest for facial point detection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 9, pp. 1507–1520, Sep. 2015.

[25] A. Telea, "An image inpainting technique based on the fast marching method," *J. Graph. Tools*, vol. 9, no. 1, pp. 23–34, Jan. 2004.

[26] F. Song, X. Tan, S. Chen, and Z.-H. Zhou, "A literature survey on robust and efficient eye localization in real-life scenarios," *Pattern Recognit.*, vol. 46, no. 12, pp. 3157–3173, Dec. 2013.

[27] R. Valenti and T. Gevers, "Accurate eye center location and tracking using isophote curvature," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1452–1459.

[28] M. Leo, D. Cazzato, T. De Marco, and C. Distante, "Unsupervised eye pupil localization through differential geometry and local self-similarity matching," *PLoS ONE*, vol. 9, no. 8, Aug. 2014, Art. no. e102829.

[29] W. Zhang, M. L. Smith, L. N. Smith, and A. Farooq, "Eye center localization and gaze gesture recognition for human–computer interaction," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 33, p. 314, 2016, doi: 10.1364/JOSAA.33.000314.

[30] A. Laddi and N. R. Prakash, "An augmented image gradients based supervised regression technique for iris center localization," *Multimedia Tools Appl.*, vol. 76, no. 5, pp. 7129–7139, Mar. 2017.

[31] H.-B. Cai, H. Yu, C.-Y. Yao, S.-Y. Chen, and H.-H. Liu, "Convolution-based means of gradient for fast eye center localization," in *Proc. Int. Conf. Mach. Learn. (ICMLC)*, vol. 2, Jul. 2015, pp. 759–764.

[32] H. Cai, B. Liu, J. Zhang, S. Chen, and H. Liu, "Visual focus of attention estimation using eye center localization," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1320–1325, Sep. 2017.

[33] G. M. Araujo, F. M. L. Ribeiro, E. A. B. Silva, and S. K. Goldenstein, "Fast eye localization without a face model using inner product detectors," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 1366–1370.

[34] B. D. Eddine, F. N. Dos Santos, B. Boulebtateche, and S. Bensaoula, "EyeLSD a robust approach for eye localization and state detection," *J. Signal Process. Syst.*, vol. 90, no. 1, pp. 99–125, Jan. 2018.

[35] N. Poulopoulos and E. Z. Psarakis, "Real time eye localization and tracking," in *Advances in Service and Industrial Robotics*. Cham, Switzerland: Springer, 2019, pp. 560–571.

[36] *Talking Face Video*. Accessed: Mar. 20, 2019. [Online]. Available: http://www-prima.inrialpes.fr/FGnet/data/01-TalkingFace/talking_face.html

[37] *The MUCT Face Database*. Accessed: Mar. 20, 2019. [Online]. Available: http://www.milbo.org/muct/

[38] Z. Pang, C. Wei, D. Teng, D. Chen, and H. Tan, "Robust eye center localization through face alignment and invariant isocentric patterns," *PLoS ONE*, vol. 10, no. 10, Oct. 2015, Art. no. e0139098.

[39] M. De Marsico, C. Galdi, M. Nappi, and D. Riccio, "FIRME: Face and iris recognition for mobile engagement," *Image Vis. Comput.*, vol. 32, no. 12, pp. 1161–1172, Dec. 2014.

[40] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Mobile iris challenge evaluation (MICHE)-I, biometric iris dataset and protocols," *Pattern Recognit. Lett.*, vol. 57, pp. 17–23, May 2015.

**JINGUO LIU** (Senior Member, IEEE) received the Ph.D. degree in mechatronics from the Shenyang Institute of Automation (SIA), Chinese Academy of Sciences (CAS), in 2007. He has been a Full Professor, since January 2011, has also been the Assistant Director of the State Key Laboratory of Robotics, since 2008, and has also been the Associate Director of the Center for Space Automation Technologies and Systems, since 2015, with the Chinese Academy of Sciences. His research interests include bio-inspired robotics and space robot. He has authored or coauthored three books, over 100 articles, and holds 50 patents in above areas.

Dr. Liu is a Senior Member of the IEEE Technical Committee on Safety, Security, and Rescue Robotics, the IEEE Technical Committee on Marine Robotics, and the Chinese Mechanical Engineering Society. He was a recipient of the T. J. TARN Best Paper Award in Robotics from the 2005 IEEE International Conference on Robotics and Biomimetics, the Best Paper Award of the Chinese Mechanical Engineering Society, in 2007, the Best Paper Nomination Award from the 2008 International Symposium on Intelligent Unmanned Systems, the Best Paper Award from the 2016 China Manned Space Academic Conference, the Outstanding Paper Award from the 2017 International Conference on Intelligent Robotics and Applications, and the Best Paper Award from the 2018 International Conference on Electrical Machines and Systems. He serves as the Associate Editor or Technical Editor of several journals, such as the IEEE/ASME Transactions on Mechatronics, IEEE Access, *Mechanical Sciences, Science China Technological Sciences, Chinese Journal of Mechanical Engineering*, and *Chinese Journal of Aeronautics*.

**ZHAOJIE JU** (Senior Member, IEEE) received the B.S. degree in automatic control and the M.S. degree in intelligent robotics from the Huazhong University of Science and Technology, China, and the Ph.D. degree in intelligent robotics from the University of Portsmouth, U.K. He held research appointments at University College London, London, U.K., before he started his independent academic position at the University of Portsmouth, in 2012. He has authored or coauthored over 170 publications in journals, book chapters, and conference proceedings. His research interests include machine intelligence, pattern recognition and their applications on human motion analysis, multifingered robotic hand control, human–robot interaction and collaboration, and robot skill learning.

Dr. Ju received four Best Paper Awards and one Best AE Award in ICRA2018. He is an Associate Editor of the IEEE Transactions on Cybernetics, *Journal of Intelligent and Fuzzy Systems*, *International Journal of Fuzzy Systems*, and *Chinese Journal of Mechanical Engineering*.

**LIHONG DAI** received the B.S. degree in automatic control and the M.S. degree in control theory and control engineering from the University of Science and Technology Liaoning, China, in 2000 and 2004, respectively. She is currently pursuing the Ph.D. degree in pattern recognition and intelligent system with the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China.

Since 2004, she has been a Teacher with the School of Electronic and Information Engineering, University of Science and Technology Liaoning. She is currently a Senior Lecturer. Her research interests include gaze tracking, computer vision, machine learning, pattern recognition, and their applications on human–robot interaction and collaboration.

**YANG GAO** (Senior Member, IEEE) received the B.Eng. degree in electrical and electronics engineering and the Ph.D. degree in artificial intelligence, control and instrumentation from Nanyang Technological University, Singapore, in 2000 and 2002, respectively. She is currently a Professor of space autonomous systems with the Surrey Space Centre, University of Surrey, Guildford, U.K. She is also the Hub Director of the Future AI and Robotics for Space (FAIR-SPACE), a U.K. National Centre of Research Excellence in Space Robotics and Autonomous Systems. She specializes in robotic vision, machine learning, and biomimetics for spacecraft GNC and mechanisms. She brings nearly 20 years of R&D experience in solving robotic system problems, and is actively involved in development of real-world space missions, such as ESA's ExoMars, Proba3, LUCE-Ice Mapper, U.K.'s MoonLITE/Moonraker, and China's Chang'E3.

• • •