Video Based Non-Cooperative Iris Segmentation

Yingzi Du*  Emrah Arslanturk

Department of Electrical and Computer Engineering
Purdue School of Engineering and Technology
Indiana University-Purdue University Indianapolis
Indianapolis, IN 46202
Email {yidu,earslant}@iupui.edu

ABSTRACT

Iris segmentation is one of the most important steps in an iris recognition system. Its accuracy can directly affect the recognition accuracy. For non-cooperative users, the obtained images often do not have good quality. Under such conditions, the iris may be deformed, out-of-focus, or motion blurred. Sometimes, images do not have a valid iris. It is very challenging to segment the iris efficiently and accurately under a non-cooperative scenario. In this paper, we proposed a novel segmentation method that uses a coarse to fine approach to extract the iris region. The preliminary result shows the proposed method is efficient and accurate.

Index Terms—biometrics, iris recognition, iris segmentation, non-cooperative iris segmentation

1. INTRODUCTION

Biometrics is a technology of identifying or verifying a person using physiological or behavioral characteristics, such as face, fingerprint, iris, palm-print, voice, gait and hand geometry etc. Biometrics is a convenient, secure and hard to spoof recognition technique, which is important for security of governmental and commercial organizations. Among them, iris recognition has been tested to be most accurate biometric identification system.

Every iris has a distinct pattern that not only differs from person to person, but is also different from left eye to right eye. Various methods have been proposed to encode the high level of distinction of the iris pattern for recognition of a person.

The typical iris recognition system is shown in Fig. 1, which includes image acquisition, iris segmentation, iris feature extraction, iris template generation, iris template matching, and iris identification. The very first step of iris image processing is to segment the iris. Iris segmentation includes pupillary boundary and limbic boundary detection, and eyelids and eyelash exclusion.

Over the past decade, many iris segmentation methods have been proposed. Daugman introduced the integrodifferential operator to detect the iris and pupil boundaries and modeled these boundaries as concentric circles. Wildes et al. introduced edge mapping and circular contour parameter voting technique to calculate the parameters of pupil and iris boundary circles. Ma et al. developed a segmentation method by approximating the pupil centroid coordinates and applying Canny edge detection and Hough transform only in iris region determined by center of the pupil. Huang et al. used edge detection and Hough transform with a combination of filter. Arvacheh et al. proposed an iris segmentation method using active contour fitting to circular pupil and limbic boundary.

These iris segmentation techniques work well for frontal iris image. These methods have assumed that pupillary and limbic boundaries could be modeled by circles. This is not always true in the frontal iris image. The eye is a three dimensional object. But we are using a 2-D camera to obtain the iris image. For non-frontal iris, it is deformed and the
circular assumptions may not hold. Therefore the segmentation process for non-frontal iris images is much more challenging. Schuckers et al. used integrodifferential operator for fitting an ellipse to the iris and pupil boundaries. Tisse et al. introduced an iris segmentation method based on a combination of integrodifferential operator and Hough transform. Daugman improves on his previous work by modeling the boundaries using an active contour where a Fourier series approximation is applied to polar images of the pupil and limbic region. Miyazawa et al. used highly flexible 10 parameter model on pupil and iris boundaries. Roy et al. proposed a segmentation method based on chain code, zigzag collaret area and support vector machine algorithms for pupil and limbic boundary detection, and Hough transform for eyelid detection. These iris segmentation techniques use a single image of an iris.

The accuracy of iris segmentation can affect the iris recognition accuracy dramatically. On one hand, including noisy information as part of iris pattern could affect the feature extraction accuracy. On the other hand, inaccurate modeling of pupil or limbic boundary can cause drastic differences on an encoded iris, which affects the matching score dramatically. Even for iris image databases containing only frontal images, it is shown that improving the pupil and limbic boundary model increases accuracy of the iris recognition system significantly.

Currently, most iris recognition systems are designed to work only with a focused and clear snapshot of a frontal iris while the subject is looking toward the vision system cooperatively. For a non-cooperative recognition system, where the subject is not aware of the presence of the camera, relying on quality of one image is not reasonable. Fig. 2 shows examples of iris images from non-cooperative users. In such a situation, using a video sequence can provide more information, which is crucial for accuracy of iris recognition. For non-cooperative users, there could be many noisy video frames with different quality of iris images. Poor quality images can affect recognition dramatically. Therefore, processing of video sequences is more challenging. It requires a robust and accurate segmentation method to extract high quality iris images. At the same time, it should be efficient enough to run in real-time. So far, little research has been done in this area. In this paper, we designed a real-time segmentation system for non-cooperative iris video images.
2. METHOD

The proposed method uses coarse pupil location to find the pupil boundary points and fits a closed loop conic by using “Direct Least Square Fitting of Ellipses”\textsuperscript{16}, which is an orientation invariant model of elliptical closed loop shape. The same conic fitting algorithm is used for iris boundary. Also, the proposed algorithm estimates the position of the iris by using the coefficients of pupil fitting on Parzen Window Classification method.

2.1. Clustering and Fast Coarse Pupil Detection

Video applications need high processing speed per frame in order to process each captured frame. Under non-cooperative situation, the obtained image may not even have an eye. This could be the result of eye blinking, eye occlusion, and eye motion. The first step is to determine if the pupil exists in the current frame, and if so, find the coarse location of the pupil. The video frames not containing valid iris information should first be removed.

To improve the coarse detection speed, we use 10 times down sampled images. By using the low intensity value and the close-to-circular shape of the pupil, rough position and radius of the pupil can be estimated quickly. For finding the dark regions of the image, a clustering algorithm is proposed. For high efficiency and accuracy, clustering via Principal Component Analysis (PCA) algorithm is applied on 10 times down sampled image with 5 clusters. It is assumed that the first frame has the same illumination characteristics with the rest of the video sequence. So, only the first frame in which the pupil exists is used for training of the clustering algorithm.

For clustering the image, K-mean clustering algorithm is proposed. Each pixel is clustered into 5 classes according to 8-neighbors’ intensity values. By using 8-neighbors, effects of noise and inconsistent illumination are decreased.

\[ J_k = \sum_{i=1}^{n_k} (x_i - m_k)^2 \]  
(2-1)

where \( m_k = \frac{1}{n_k} \sum_{i \in C_k} x_i \) is the centroid of class \( C_k \) and \( n_k \) is the number of elements in class \( C_k \).

\[ C_k = \arg \min_{C_k} J_k, \quad K = 1, \ldots, N \]  
(2-2)

where \( N \) is the number of classes. After a pixel is classified to a cluster, the cluster should be updated.

\[ m_k = \frac{(n_k + x) / (n_k + 1)} \]  
(2-3)

PCA is an orthogonal linear transformation of given data to a new coordinate system according to variance of the projection of the data. PCA is used to reduce the dimensions of the cluster by picking up the dimensions with the largest variances, which means finding the best low rank approximation of the data by minimizing the least squares\textsuperscript{17}. PCA, via singular value decomposition (SVD), guarantees the best lower dimensional linear approximation of the data set\textsuperscript{18}.

\[ \overline{m} = \frac{1}{N} \sum_{k=1}^{N} m_k \]  
(2-4)

where \( \overline{m} \) is the mean of centers of clusters. The sample mean is a zero-dimensional representation of the clustered data. By projecting the data onto a line passing through the sample mean, we obtain:

\[ x_k = \overline{m} + a_k e \]  
(2-5)

where \( e \) is a unit vector in the direction of the line and \( a_k \) is scalar corresponding to the distance of \( x_k \) to \( \overline{m} \). By minimizing the squared-error, optimal set of \( a_k \) can be found.

\[ J_2(a_1, \ldots, a_N) = \sum_{k=1}^{N} \| (\overline{m} + a_k e) - x_k \|^2 \]  
(2-6)

\[ J_2 = \sum_{k=1}^{N} a_k^2 \| e \|^2 - 2 \sum_{k=1}^{N} a_k e^T (x_k - \overline{m}) + \sum_{k=1}^{N} \| x_k - \overline{m} \|^2 \]  
(2-7)

The norm of unit vector \( e \) is equal to 1. The partial derivative of Eq. 2- with respect to \( a_k \) is,

\[ \frac{\partial J_2}{\partial a_k} = 2a_k - 2e^T (x_k - \overline{m}) \]  
(2-8)
For minimizing the least squared error, the derivative of the error criterion function should be zero. When the Eq. 2-8 is set to zero, we obtain,

$$a_k = e^T (x_k - \overline{m})$$  \hspace{1cm} (2-9)

Eq. 2-9 is the least squares solution of $x_k$ projected onto the line passing through the sample mean in the direction of $e$. For finding the optimum direction $e$, scatter matrix $S$ is defined as,

$$S = \sum_{k=1}^{N} (x_k - \overline{m})(x_k - \overline{m})^T$$  \hspace{1cm} (2-10)

By substituting Eqs. 2-9 and 2-10 into Eq. 2-7, we can rewrite the squared error criterion function as,  

$$f_L(e) = -e^T Se + \sum_{k=1}^{N} ||x_k - \overline{m}||^2$$  \hspace{1cm} (2-11)

By using Lagrange multipliers method, $e^T Se$ can be maximized subject to the constraint $||e|| = 1$,

$$u = e^T Se - \lambda (e^T e - 1)$$  \hspace{1cm} (2-12)

where $\lambda$ is the undetermined Lagrange multiplier. After partially differentiating Eq. 2-12 with respect to $e$, and setting the derivative to zero, we obtain,

$$Se = \lambda e$$  \hspace{1cm} (2-13)

In order to maximize $e^T Se$, the eigenvectors corresponding to the largest eigenvalues of the scatter matrix should be selected. In this way, the best one-dimensional projection of the data in least-sum-of-squared-error sense can be found. Then the $d'$ dimensional projection of the clustered data can be rewritten as,

$$x = \overline{m} + \sum_{l=1}^{d'} a_l e_l$$  \hspace{1cm} (2-14)

where $e_l$ for $l = 1, ..., d'$ are the $d'$ eigenvectors of the scatter matrix corresponding to the largest eigenvalues, and the coefficients $a_l$ for $l = 1, ..., d'$ are the principal components of $x$ in the basis of $e$.

The principal components and basis vectors calculation is only applied on the first frame that a pupil exists. For clustering the following frames of the video sequence, previously calculated mean of clustered data ($\overline{m}$), principal components ($a_l$) and basis vectors of projected eigen-space ($e_l$) are used.

![Figure 3. Clustered Iris Frames Captured from Same Video Sequence.](image-url)

After clustering a frame, Fig. 3, finding the coarse pupil location and radius is trivial. Each frame of a video sequence is in an order, so the pupil location cannot change dramatically in two consecutive frames. By using this assumption, if the pupil is found in the previous frame the boundary is found around the previous pupil center. However, if the pupil could not be found in the previous frame or the current frame is the first frame, then the course pupil location and radius are found by using convolution with edge information found from clustered image.
2.2. Pupil and Iris Boundary Detection via Fast Ellipse Fitting

By using the found coarse pupil location and radius, the region of interest (ROI) is defined around the rough pupil location. The motion of the eye changes the orientation of the pupil with respect to camera axis. Regular circle or ellipse fitting algorithms are insufficient to model the boundaries of a moving pupil. Ellipse can be formalized in Cartesian coordinates as:

\[
\frac{x^2}{r_1^2} + \frac{y^2}{r_2^2} = 1
\]  
(2-15)

where \( r_1 \) and \( r_2 \) are radius corresponding to \( x \) and \( y \) coordinates respectively. When \( r_1 \) and \( r_2 \) are equal, the equation represents a circle. However, this equation does not model rotation of an ellipse. For frontal iris images, the pupil can be modeled as a perfect circle or ellipse because of almost zero rotation of the pupil according to camera axis. On the other hand, with non-cooperative iris images this assumption cannot be used. Each position of the iris except the center position creates a rotation of the pupil. For modeling the moving eye, a more complex elliptical model is needed. “Direct Least Square Fitting of Ellipses” method is proposed to mathematically model the pupil boundary. The advantages of this algorithm are: ellipse-specificity, providing useful results under occlusion; invariance to affine transformation; high robustness to noise; and high computational efficiency.

General conic can be represented by second order polynomial;

\[
F(a, x) = a \cdot x = a x^2 + b x y + c y^2 + d x + e y + f = 0
\]  
(2-16)

where \( a = [a b c d e f]^T \) and \( x = [x^2 x y y^2 x y 1]^T \). This model is any 2 dimensional elliptical section of a 3 dimensional conic. The \( x \) and \( y \) values are the pixel coordinates of the edge points of the pupil boundary. However, second order polynomial representation assumes the coordinates are in Cartesian coordinates. The pupil boundary coordinates should be normalized before fitting.

A simple constraint applied on the linear least squares problem provides the high efficiency of the ellipse fitting algorithm. The constraint is the discriminant of the roots of the second order polynomial should be equal to 1.

\[
b^2 - 4ac = 1
\]  
(2-17)

This quadratic constraint can be expressed in matrix form as \( a^T C a = 1 \).

\[
C = \begin{bmatrix}
0 & 0 & 2 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]  
(2-18)

Now the constrained ellipse fitting problem reduces to

\[
\text{minimizing } E = \|D a\|^2 \text{ subject to constraint } a^T C a = 1
\]  
(2-19)

where the design matrix \( D \) is defined as \( D = [x_1 x_2 \ldots x_N]^T \).

The direct ellipse-specific fitting algorithm assumes the input is normalized, so the estimated coefficients found from minimizing the least squares are normalized coefficients. After finding the actual coefficients, normalized coefficients will be used to estimate pupil location.
The advantage of using the proposed method is the robustness of the ellipse fitting algorithm against noisy data and occlusion, which are two common elements of pupil detection. By using this method, the effects of occlusion of the pupil with eyelids, and the noisy information of eyelashes and glare are minimized. Beside the robustness, the efficiency of the algorithm makes it ideal for video based segmentation. The ellipse fitting algorithm runs under 0.4 milliseconds using Matlab.

A similar approach is used for iris detection. The iris boundary is almost a perfect circle when the eye is looking to the front. Because there is no muscle controlling the shape of the iris boundary, the radius of the iris boundary only changes with the distance of the camera to the eye. However, for the non-cooperative video frames, the shape of the iris also changes with the motion of the eye. Even if it is almost a perfect circle at frontal gaze, the iris boundary has elliptical shape when the eye is looking another direction. The rotation of the eye corresponding to the camera axis creates affine transformation of the iris boundary. The same ellipse fitting algorithm is used for modeling the iris boundary. However, finding the iris boundary points is more difficult than for the pupil boundary. The pupil boundary has stronger edges than the iris boundary, and the occlusion of the eyelids is less for the pupil boundary when compared with the iris boundary. An adaptive threshold and edge detection algorithm is used to determine the edge points that will be used in ellipse fitting.

As seen in Fig. 5, the robustness of the ellipse fitting algorithm against the occlusion helps iris boundary fitting against the eyelids.

One of the key steps of iris recognition is segmentation of the iris region from the rest of the image accurately. Locating the eye and extracting the pupil and iris boundaries in the image is not trivial for non-cooperative iris video frames. The location and orientation of the eye and occlusion of the eyelids and eyelashes creates non-ideal circumstances for segmentation. By using the robustness of the “Direct Least Square Ellipse Fitting” algorithm, the negative effects of occluded and noisy data is minimized and the efficiency of the algorithm makes the proposed method crucial for video based image processing.
2.3. Eyelid Detection

The iris area generally is occluded by the eyelids, and before the identification process, the occluded part should be eliminated. For frontal gaze images, it can be assumed that eyelids are the horizontally strongest edges. Also, there is eyelid information on both sides of the iris area. However, for non-cooperative iris images, where the eye can look to any direction, the shape of the eyelid is changing. The disadvantage of this change can cause mixing of the pupil and iris boundary with eyelid information. Moreover, the assumption of having eyelid information on both sides of the iris area is not always true.

For eyelid detection on non-cooperative iris images, it is proposed to use only the edge information between the pupil area and the iris area. By this limitation, the eyelid information outside the iris area will be eliminated with the noise.

By modeling top and bottom eyelids separately with a second degree parabola, the occluded parts can be removed from the iris area. The preliminary results show that for most cases the eyelids can be detected and modeled with the proposed method.

Figure 6. Output of Segmentation Algorithm.

However, with the severe occlusion cases the detection results are poor and need to be optimized. In cases of eyelids covering the pupil area, accuracy of segmentation of the eyelid is decreased.

3. EXPERIMENTAL RESULTS

In our experiments, our own database is used to test the segmentation algorithm. The database is taken in our laboratory with a consumer camcorder (SONY DCR-SR100) and near infrared light source used for the night shot (SONY HVL-IRM). Because the setup is not fixed, and the light sources are powered with battery, the video sequences have inconsistent illumination, variable distance between subject and the camera, and variable pupil and iris radii. That is why it can be claimed that our database is suitable for experimenting using our proposed non-cooperative iris segmentation method.
Our database consists of 27 subjects, 54 eyes (left and right eye of each subject) and 186 video sequences, 59,648 frames. The proposed method is run on each video sequence separately and the bad frames are eliminated automatically. Some bad frames, like a frame when the subject is blinking or the eyelid is occluding the iris severely, are eliminated during the segmentation process. However, the elimination of some bad frames is not that trivial. Some images with motion blur, out of focus or inconsistent or poor illumination could not be eliminated, and the result of the segmentation can be poor.

As seen in Fig. 8, the images in our database were not taken under ideal conditions. Besides being hard or impossible to segment the iris, these images are also not suitable for any kind of recognition algorithm. After processing each video, 9 videos out of 186 did not have enough segmented frames for recognition. The most common reason of poor segmentation of those videos is very poor illumination, where even the pupil and iris boundary cannot be separated. Another reason is poor focus of the camera, where there is minimum or no strong edge information of pupil and iris.
boundary. However, even if most of those videos are segmented properly, the information of the iris patterns is not available for encoding and matching.

4. CONCLUSION

In this paper, a video based non-cooperative iris segmentation method is proposed. The aim of the method is to accurately extract the iris region from the image and clear the occluded and noisy data. To improve the efficiency of the method, clustering and a coarse pupil detection algorithm is applied, where the frames containing no pupil are also eliminated. For modeling the pupil and iris boundary, “Direct Least Square Fitting of an Ellipse” algorithm is used. The applied ellipse fitting algorithm is invariant against affine transformation, efficient, and robust against occlusion and noise. The experimental results of pupil and iris segmentation of non-cooperative iris frames are accurate. In order to remove the eyelids, a second order parabola is fitted to the detected edge information between the pupil and iris boundaries. The results of the eyelid removal algorithm is promising, however, it needs to be improved for severely occluded and poorly illuminated iris frames. The proposed iris segmentation algorithm aims to segment an iris from a frame in non-ideal conditions, which would lead to implementation of an iris recognition system for surveillance.

ACKNOWLEDGEMENT

This project is sponsored by the ONR Young Investigator Program (award number: N00014-07-1-0788) and National Institute of Justice (award number: 2007-DE-BX-K182). The data of this project is partially collected by C. Belcher and Z. Zhou in our Pattern Recognition and Biometrics Laboratory. We would also like to thank the people who contributed their iris data for this project.

REFERENCES


