

An Automated Video-Based System for Iris Recognition

Yooyoung Lee, P. Jonathon Phillips, and Ross J. Micheals

National Institute of Standards and Technology, Information Technology Laboratory
100 Bureau Drive, Gaithersburg, MD 20899-8940
{yooyoung, jonathon, rossm}@nist.gov

Abstract. We have successfully implemented a Video-based Automated System for Iris Recognition (VASIR), evaluating its successful performance on the MBGC dataset. The proposed method facilitates the ultimate goal of automatically detecting an eye area, extracting eye images, and selecting the best quality iris image from video frames. The selection method's performance is evaluated by comparing it to the selection performed by humans. Masek's algorithm was adapted to segment and normalize the iris region. Encoding the iris pattern and then completing the matching followed this stage. The iris templates from video images were compared to pre-existing still iris images for the purpose of the verification. This experiment has shown that even under varying illumination conditions, low quality, and off-angle video imagery, that iris recognition is feasible. Furthermore, our study showed that in practice an automated best image selection is nearly equivalent to human selection.

Keywords: Biometrics, Iris recognition, Eye detection, Image quality measurement, VASIR, Iris segmentation.

1 Introduction

Biometrics is the umbrella term for methods that recognize an individual based on physiological or behavioral characteristics. The human iris is a highly distinctive feature of an individual to establish identity of an individual with very high accuracy.

Still-image to still-image comparison for iris recognition is routinely covered by existing research studies. However, video-image to still-image comparison is a relatively new research subject which needs to overcome a number of challenges before the discipline can receive widespread acceptance. Challenges include recognizing a person in infrared image sequences, coping with high and low resolution, processing video sequences of people walking through a portal, matching to still face images, etc. The U.S. government sponsored the Multiple Biometric Grand Challenge (MBGC) program, which provides a standardized dataset as an aid to find solutions and to advance the current state of iris recognition technology[1].

Video-based Automated System for Iris Recognition (VASIR) is a method that successfully overcomes current limitations caused by external influences such as inconsistent lighting conditions, or low image quality. Its performance and practical feasibility have been evaluated using the above-mentioned MBGC dataset.

This paper also proposes a method for the selection of the best quality iris image from video frames using an automated algorithm. This method is a confirmation of its viability when compared against manual human selection.

This paper is divided into seven sections. This section served to introduce the notion of automated system for iris recognition in video streams. Section 2 gives an overview on the VASIR procedure. Sections 3-5 describe the eye detection/ extraction, image quality measurement, and the iris recognition algorithm. The experimental results are illustrated in section 6. Section 7 contains results and makes suggestions for future work.

2 Overview

Generally, still-image to still-image iris recognition proceeds in four distinct steps. The first is the image acquisition, which yields an image of the subject's eye region. The second step is the iris localization/normalization, which segments the iris from the rest of the acquired image. As a third step the iris image is encoded in binary. The last step compares an existing iris pattern (gallery) with the generated iris pattern (probe), to provide a decision on whether two irises are the same.

On the other hand, matching a video-based template to a still-based template needs more components for its iris verification procedure. Fig.1 illustrates the basic function of the VASIR process.

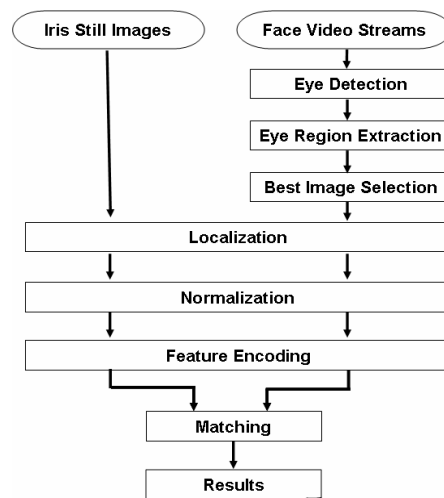


Fig. 1. VASIR Procedure

Before the localization (previously step 2), three additional steps were added in our implementation of the iris verification procedure. The new initial step is the automatic detection of the eye region within video frames; followed by the eye region extraction. In the last additional step, we select the “best” quality image from video.

3 Eye Detection and Extraction

The term “eye detection” will be used to refer to the process of the continuous eye area detection in video streams. We employed a method called “object detector”, initially proposed by Viola-Johns[2] and later improved by Lienhart-Maydt[3] through the addition of rotated Haar-like features as described below. The integral image, an intermediate image representation, utilized in the method allows a rapid computation of features. The yielded features are then filtered by classifiers which are based on the boost learning algorithm. Combining multiple classifiers in a cascade allows for the background regions of the image to be quickly rejected, enabling more computation on promising eye pair like regions. A cascade of classifiers has a decision tree; each node is trained to identify objects-of-interest and to reject unrelated objects. In combination with a search window moving across the frame, the classifier can be used to locate the eye region. The classifier is resizable in order to be able to find objects-of-interest independent of varying size, i.e. the scan procedure needs to be repeated several times to find objects of differing size.

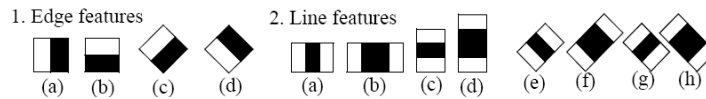


Fig. 2. Used Haar-like Features (captured from [3])

The edge features measure the difference in intensity between the eye regions and the region across the upper cheeks. The features make use of the fact that the eye region is often darker than the cheeks. The line features compare the intensities in the eye regions to the intensity across the bridge of the nose.

The OpenCV community[4] shared a collection of public domain classifiers and the performance of these classifiers were evaluated by On-Santana et al[5]. To detect both the left and right eye region – also called eye-pair or two-eye – in a face video stream, we chose the named eye-pair classifier[6][7] due to its fitting properties, speed and size; The eye-pair classifier is applied to each frame to find the position and size of both eye regions. The classifier in the cascade is 45x11 pixels in size and consists of 19 stages.

We defined a minimum size for the eye area in the frames to lower the number of false positives. The detected region is saved into three separate files: two-eye, left-eye-only and right-eye-only automatically for analysis purposes. Detailed performance results are discussed in section 6.

4 Image Quality Measurement

Normally, present commercial iris recognition systems use images acquired under constrained conditions on account of their image quality requirements. Near-infrared illumination is used to light the face, and the user is prompted with visual and/or auditory feedback to correct the eye position. Due to this the image is often in focus and of sufficient size for iris recognition.

Being able to measure the eye image quality in video streams under general unconstrained conditions would permit a more accurate determination when to match two iris templates. Several groups[8] have studied how to determine the focus of an image (e.g. analyzing the sharpness of the pupil and iris boundary, calculating the total high frequency power in the 2D Fourier spectrum, a variation of the sum modulus difference (SMD) filter, etc).

This paper proposes two methods of selecting the best image from eye images obtained using video streams. The first method is named the “Human Image Quality Measurement” (HIQM), where the best quality eye images are manually selected via human vision. The other method is an application of the “Automatic Image Quality Measurement Algorithm” (AIQM) via the edge density, which has been proposed for predicting the face image quality by Beveridge et al[9]. Fig.3 illustrates the processes.

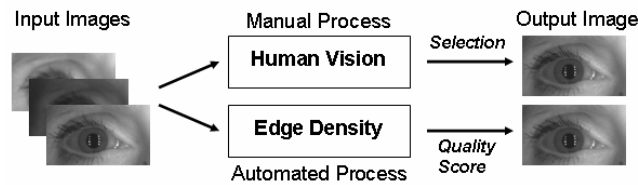


Fig. 3. Image Quality Measurement (via Human Vision and Edge Density)

Out of the set of eye images, and for each eye individually, two images are selected by human vision in HIQM. Afterwards, the two selected eye images are segmented to isolate the iris region, and the best one-eye images (left, right) are chosen corresponding to segmentation results. The detailed segmentation process will be addressed in section 5.

For AIQM, the average edge density in an image is calculated, to determine the level of focus. The sobel edge detector is applied to the image to derive the edges and the average sobel edge is quantified within the size of the eye region.

The performance of VASIR is evaluated as to whether the AIQM is equivalent to HIQM for best image selection. Performance values will be provided in section 6.

5 Iris Recognition Algorithm

The iris verification components in VASIR were implemented using the algorithm initially developed by Masek[10], re-implemented for irisBEE[11].

In the *iris localization* step, the iris is segmented by detecting the inner circle, which is the boundary between the pupil and the iris, and the outer circle, which is the boundary between the iris and the sclera. An automatic segmentation algorithm based on the circular Hough Transform initially employed by Wildes et al[12] is used. The center coordinates (x_c, y_c) and the radius r are able to define any circle according to the equation;

$$x_c^2 + y_c^2 - r^2 = 0 \quad (1)$$

Masek's algorithm uses circular Hough Transform to detect the iris and pupil boundaries by employing Canny edge [10] to generate the edge map. The eyelids were removed using a linear Hough Transform by inserting horizontal lines attached to the eyelids itself. The eyelashes are detected and ignored using threshold manipulation. VASIR applies the segmentation to extract the iris region from the eye images selected by the image quality measurement algorithm.

For *normalization* step, the outcome is a standardized 2D representation of the iris pattern, regardless of the pupil dilation, source image size, non-concentric pupil displacement, or the distance of the eye to the capture device. Daugman[13]'s polar coordinate-based system effectively eliminates scaling issues caused by linear stretching.

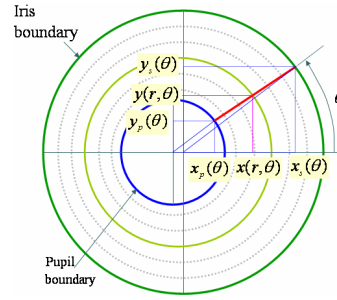


Fig. 4. Doubly Dimensionless Projected Polar Coordinate System¹

The homogeneous rubber sheet model assigns to each point within the iris region a pair of real coordinates (r, θ) where r lies on the unit interval $[0,1]$ and θ is the angle over $[0,2\pi]$. The remapping of the iris image $I(x,y)$ from Cartesian coordinate (x,y) to polar coordinate (r, θ) is classically represented as;

$$\begin{aligned} I(x(r, \theta), y(r, \theta)) &\rightarrow I(r, \theta) \\ x(r, \theta) &= (1-r)x_p(\theta) + rx_s(\theta) \\ y(r, \theta) &= (1-r)y_p(\theta) + ry_s(\theta) \end{aligned} \quad (2)$$

$(x_p(\theta), y_p(\theta))$ and $(x_s(\theta), y_s(\theta))$ are the coordinates of the pupil and iris boundaries along the θ direction. In Masek's algorithm, the pupil center is used as reference point, and the iris region is scaled based on the radial vectors. Finally, an iris pattern image and the noise mask using occlusion information produced in the localization stage are generated containing the size of the angle and the radial.

Many different filters have been suggested to extract the features from the iris region (e.g. Laplacian-of-Gaussian filter, Wavelet Transform, Discrete Cosine Transform, etc.) [8]. Masek's algorithm employed a 1D Log-Gabor filter which was introduced by Yao et al[14], in order to process the *feature encoding* from normalized iris images. The normalized iris patterns are then coded by filters that have Gaussian

¹ The figure was modified from the presentation "Analysis of Daugman's High Confidence Visual Recognition of Persons by a Test of Statistical Independence" by Minhwan Kim.

transfer functions when viewed on the logarithmic frequency scale. The frequency response of a Log-Gabor filter is given as;

$$G(w) = e^{(-\log(w/w_0))^2 / (2\log(\sigma/w_0))^2} \quad (4)$$

Where w_0 represents the filter's centre frequency, and σ gives the filter's bandwidth. Each isolated iris pattern is demodulated to extract its phase information using Daugman's method[15].

For *matching*, Masek employed the Hamming Distance (HD) that incorporates noise masking. The HD measure can be used to make a decision whether the iris pattern is of the same person or a different person. The noise mask helps to use only the significant bits in calculating the HD between two iris templates. For rotational inconsistencies between two iris templates, one template is shifted left and right bit-wise and selects values from successive shifts[15]. If an angular resolution of 180 is used, each shift will correspond to a rotation of 2 degrees in the iris region.

6 Experimental Results

6.1 Dataset

VASIR is evaluated using the datasets collected by the MBGC[1]. Two of the goals of the MBGC are to promote and to advance iris recognition technology. The MBGC dataset includes eye images on varying illumination conditions, low quality, and off-angle or occluded images on both still and video imagery.

This paper uses two types of the datasets which we will call 'face visible videos' and 'iris still images'. The face visible videos were captured with a camera by Sarnoff with the Iris On the Move (IOM) system; each frame is 2048x2048 pixels in resolution. An LG2200 EOU camera was used to capture the still images in 640x480 resolutions. There are a total of 149 video sequences and a total of 1,668 iris still images. Of the 149, there is no eye visible in just one of the video sequences, and, in 18 other video sequences only one eye is visible. Since our purpose is to detect both left and right eyes (two-eye) within video frames, we conducted our experiment using only the 130 videos where both eyes are visible.

6.2 Performance for Eye Detection and Extraction

The number of successful two-eye detections in videos was 127 out of the total of 130 videos; in 97.69% both eyes were correctly detected. By defining a minimum size for the eye pair region, the false positives were reduced, resulting in just 3.15%.

Fig.5 depicts examples of successful two-eye detections and false positive detections. For analysis purposes, VASIR automatically extracted 1,083 left-eye and 1,083 right-eye images out of the total of 2,534 frames from the 130 video sequences. The extracted eye images were used to determine the best quality image; the selection is explained in the next section. The sizes of the extracted eye regions are very diverse due to the fact that the subject keeps moving in the video – effectively changing the distance from the camera. The average dimension of the saved images for the left and right is about 500 (width) x ¾ of the width (height) pixels in resolution.

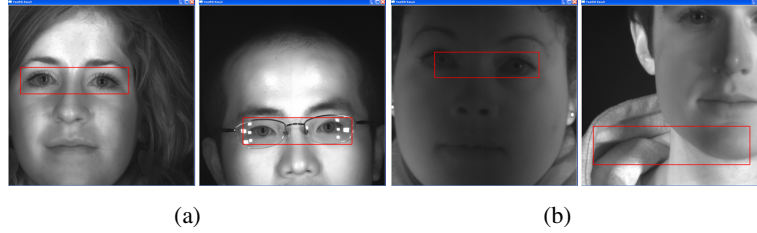


Fig. 5. Examples of the Eye Detection (a) Successful Two-Eye Detection (b) False Positive Two-Eye Detection

6.3 Best Iris Image Selection and Segmentation

As stated in section 4, the best quality image was selected by two methods: using the HIQM and through the AIQM. Fig. 6 shows an example of the original video frames (6a), the extracted left eyes (6b), and the best-left-eye image selection by AIQM.

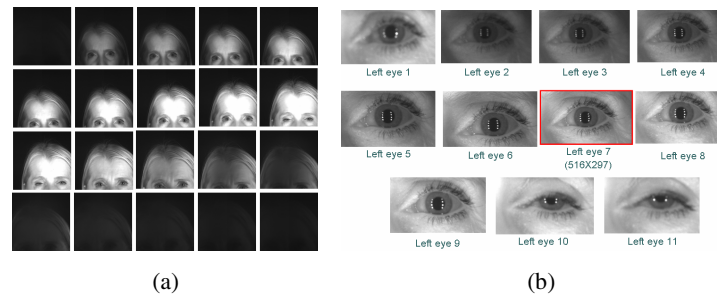


Fig. 6. An Illustration of the Best Image Selection Process (a) Original Video Frames (b) Best-Left-Eye Selected by AIQM

In the Fig. 6-b, 11 left eyes were extracted in a total 20 frames, and then the seventh left-eye was selected as the best iris image through Edge Density algorithm.

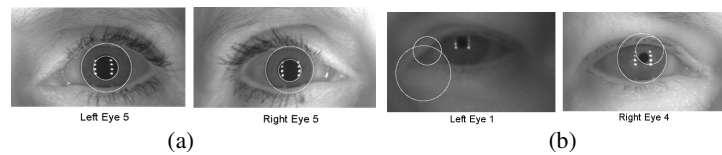


Fig. 7. Examples of the Iris Region Segmentation (a) Successful Segmentations (b) Incorrect Segmentations

All selected images were segmented in order to extract the iris region. Fig. 7 shows examples of successful and incorrect segmentations.

In the 254 (127 x 2 eyes/image) eye images, the segmentation rate for HIQM was 81.5% and 74.8% for AIQM; it shows the near-equivalency of AIQM and HIQM.

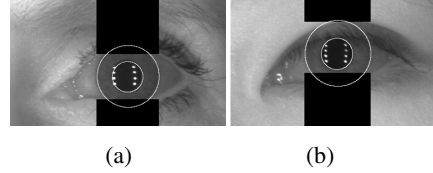


Fig. 8. Examples of the Occlusion Deletion from Eyelids and Eyelash (a) Succeed to detect occlusions (b) Failed to detect occlusions

Fig. 8 illustrates some examples of the occlusion detection to remove eyelids, eye-lashes, and reflection. However, there has been a high rate of false occlusion detection (over 45%) which influenced the failed matching results.

6.4 Normalization

The iris region is rescaled in the normalization process, to work with fixed dimension. Plus, noise masking filters out bits insignificant to the matching process. Fig. 9 shows an illustration of the normalization.

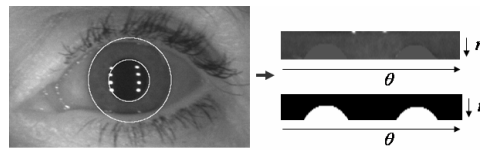


Fig. 9. Normalization Result with the Noise Mask

6.5 Matching Result for Verification (1:1)

For 4 video-images out of the total 127 videos, there were no corresponding still-images. We thus used a total of 123 video-images and the corresponding still-images; 25 persons were featured twice in the 123 datasets. There were either 12 or 6 left and right images per each person in the still-image datasets. In terms of verification, we compared each iris template from video-images with all eye images from the same person. We chose the still-image closest to the video-image according to the computed match score (Hamming Distance, HD). Each and all of the 123 video-images are compared to the 123 still-images – excluding images of the same person, to compute the non-match score. Based on the Match and Non-Match HD distributions, we can choose a separate point which allows a decision to be made when comparing two templates. However, the Match and Non-match distributions may overlap, which would result in a number of incorrect matches or false accepts, and a number of mismatches or false rejects.

The False Accept Rates (FAR) and False Reject Rates (FRR) can be calculated by the overlap area between the two distributions. The Fig. 10 illustrates the Match and Non-match HD distributions with overlap. The generated template is 240x20 pixels in resolution and 10 bits in the template were shifted left and right for rotational inconsistencies. When using a Hamming Distance value of 0.39 for the separation point,

the FAR is 0.80% and the FRR is 43.90% in HIQM; the FAR is 0.83% and the FRR is 45.53% in AIQM. Given this, if two iris templates have a Hamming Distance lower than the separation point then it is concluded that the iris templates belong to the same person; otherwise they are from a different person.

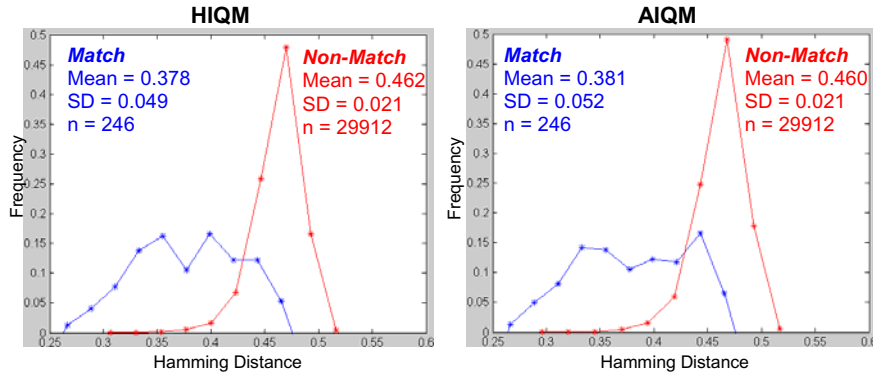


Fig. 10. Match and Non-match Hamming Distance Distributions with Overlap

The experiment resulted in a corresponding matching rate between still-images and video-images of 56.09% in HIQM and 54.47% in AIQM. Comparing HIQM with AIQM, this experiment confirmed their near-equivalency. Regarding system feasibility, the matching rate is relatively low, but system improvements are in development.

7 Conclusion

We implemented the Video-based Automated System for Iris Recognition (VASIR) and evaluated its performance on the MBGC datasets. VASIR used algorithms employing from the OpenCV library for two-eye detection, adapting human and edge density methods for the best image selection, and applying Masek's algorithm for iris recognition. A cross-platform Qt toolkit used to implement the graphical user interface. VASIR successful rate was 97.69% in two-eye detection, 81.50% in the iris region segmentation. The matching rate between a still-image template and video-image template is 56.09% in HIQM and 54.47% in AIQM.

We have shown that even under varying illumination conditions, low quality, and off-angle on both still and video imagery, iris recognition is nevertheless feasible. Furthermore, our study showed that in practice an automated best image selection is nearly equivalent to human selection.

However, VASIR still has opportunities for further improvement. There are a number of VASIR enhancements underway. We are working on new methods to improve each step as shown in Fig.1. In particular, the localization needs to be improved, by substituting the occlusion detection algorithm. Furthermore, enhancing the image quality before the localization should prevent incorrect segmentation of the iris region, caused by false focus, wrong brightness level, and moving objects.

Acknowledgments. The Authors would like to thank James J. Filliben who was extremely helpful in analyzing datasets and plotting the graphs. The identification of any commercial product or trade name does not imply endorsement or recommendation by the National Institute of Standards and Technology.

References

1. Multiple Biometric Grand Challenge, <http://face.nist.gov/mbgc/>
2. Viola, P., Jones, M.J.: Robust real-time face detection. *International Journal of Computer Vision* 57(2), 151–173 (2004)
3. Lienhart, R., Maydt, J.: An Extended Set of Haar-like Features for Rapid Object Detection. *IEEE International Conference on Image Processing*, vol. 1, pp. 900–903 (2002)
4. Intel Open Source Computer Vision Library, v1.0 (2006), <http://sourceforge.net/projects/opencvlibrary/>
5. Castrillón-Santana, M., Déeniz-Suárez, O., Antón-Canalís, L., Lorenzo-Navarro, J.: Face and Facial Feature Detection Evaluation. In: *International Conference on Computer Vision Theory and Applications* (2008)
6. Castrillón-Santana, M., Déeniz-Suárez, O., Tejera, M.H., Artal, C.G.: ENCARA2: Real-time detection of multiple faces at different resolutions in video streams. *Journal of Visual Communication and Image Representation*, 130–140 (2007)
7. Reimondo, A.: Haar cascades repository (2007), <http://alereimondo.no-ip.org/OpenCV/34>
8. Bowyer, K.W., Hollingsworth, K., Flynn, P.J.: Image Understanding for Iris Biometrics: a Survey. *Computer Vision and Image Understanding* 110(2), 281–307 (2008)
9. Beveridge, J.R., Givens, G.H., Phillips, P.J., Draper, B.A., Yui Man Lui, C.: Focus on Quality, Predicting FRVT 2006 Performance. In: *8th IEEE International Conference on Automatic Face and Gesture Recognition, FG* (2008)
10. Masek, L.: Recognition of Human Iris Patterns for Biometric Identification, The University of Western Australia, <http://www.csse.uwa.edu.au/~pk/studentprojects/libor/>
11. Phillips, J., Bowyer, K., Flynn, P., Liu, X., Scruggs, T.: The Iris Challenge Evaluation 2005. In: *IEEE 2nd International Conference on Biometrics Theory, Applications and Systems* (2008)
12. Richard, P., Wildes, C.: Iris Recognition: An Emerging Biometric Technology. *Proceedings of the IEEE* 85(9) (1997)
13. Daugman, J.G.: High Confidence Visual Recognition of Persons by a Test of statistical Independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15(11) (1993)
14. Yao, P., Li, J., Ye, X., Zhuang, Z., Li, B.: Iris Recognition Algorithm using Modified Log-Gabor Filters. In: *Proceedings of International Conference on Pattern Recognition*, pp. 461–464 (2006)
15. Daugman, J.G.: How iris recognition works. In: *Proceedings of International Conference on Image Processing*, vol. 1 (2002)