Automatic Iris Mask Refinement for High Performance Iris Recognition

Yung-hui Li and Marios Savvides
Carnegie Mellon University

***Abstract*—How to estimate artifacts on an iris image in polar domain is an important question for any iris recognition system which pursues high recognition rate as its goal. In literature, there are many different existing algorithm that estimate iris occlusion in either Cartesian or polar coordinate. In this paper, our goal is not to propose another new method to compete with existing method. Rather, our goal is to propose a new algorithm which can take any iris mask estimated by existing algorithm, and refine it into a much more accurate mask. In this way, our proposed method could co-work with any other existing algorithm and improve iris recognition performance. Experimental results show our proposed method can improve iris recognition rate by a great lead compared to the performance of the system using the unrefined iris masks.**

***Index Terms*—biometrics, iris recognition, iris mask estimation, mutual information**

I. Introduction

IRIS recognition has become one of the most accurate modalities in biometrics. For nearly two decades, researches have been mostly focused on how to segment iris more accurately and what are the best features to use for iris recognition. While these two areas still remain their importance in iris recognition, there is one link between these two stages which is usually missing or neglected in literature. This missing link is how to generate iris mask accurately and automatically.

In iris recognition, usually we will unwrap the annular-shaped iris region into a rectangle map in polar coordinate, where horizontal axis represents angular variation and vertical axis represents radius variation. This step is called iris normalization because it normalizes the dilation and contraction of pupillary motion. It also normalizes the size of iris since iris size may vary across different image acquisition devices. Just like in face recognition, where face images have to be resized and aligned before registration, if we would like to perform iris recognition in original Cartesian coordinate, how to register irises of different size becomes a big problem too. Another benefit of iris normalization is that it transforms the rotational variation of iris in Cartesian coordinate into a horizontal translation in polar coordinate, which simplifies the matching process too.

Iris images, in polar coordinate, contain two kinds of regions. The first region is the authentic iris texture, and the second region is the artifacts that occludes the authentic iris texture. The second kind of region includes eyelids, eyelashes, shadows caused by eyelid or eyelashes, and specular reflections. All of these artifacts have to be masked out and discarded during the iris matching stage. If those artifacts are not masked out, or the mask is not accurate enough to indicate all occluded regions, part of the artifacts would get into the matching stage and the iris matching performance will be affected substantially. Figure 1 shows example images of an iris texture map and its accurate mask.

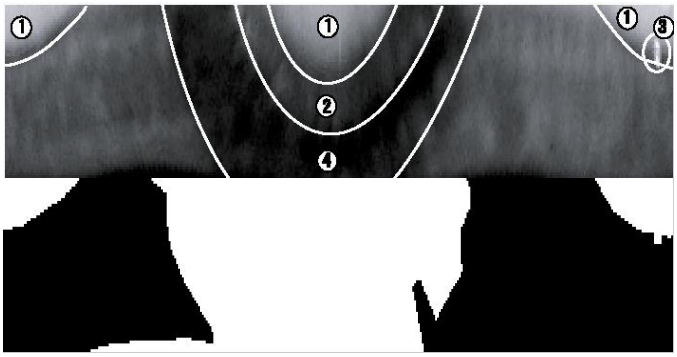


Figure 1. Normalized iris texture map (upper picture) and its accurate mask (lower picture), with white color indicates occluded area. In this iris map, there are noises caused by (1) eyelids, (2) eyelashes, (3) specular reflections, and (4) shadows, as indicated in the picture. All of these artifacts have to be indicated in the mask in order to achieve high recognition performance.

There are two approaches to estimate accurate iris mask from images. The first is to estimate the occluded region of original images in Cartesian coordinates and then transform the mask into polar coordinate. The second is to estimate the occlusion from the image in polar domain directly. In this paper, the method we proposed belongs to the second category. The advantage of estimating occluded region in polar domain is similar to the benefit we can get by doing iris normalization. Because in Cartesian domain, there are so many factors that varied all the time and are very difficult to model, modeling occluded region in polar domain simplifies the problem and also increases the computational efficiency.

We proposed a learning-based method which refines existing estimations of the occluded region on iris images in polar domain, by maximizing the mutual information on bit-level. This method is simple, easy to implement and computational efficient. Experiments also show the mask refined by proposed method can enhance the recognition performance, compared with original masks. We will review the previous work in Section II and describe our proposed method in Section III. Experimental settings and results are illustrated in Section IV. Discussion and Conclusions are presented in Section V. Finally, we point out possible future works in Section VI.

II. Previous Work

In [1], Daugman proposed an optimization scheme for finding the spline parameters that best describe the eyelid boundary. In his later work [2], he proposed a new method, which used active contour (Snake) to find the boundary of the eyelids. On the other hand, in [3], Zhang et al. proposed using Sobel edge filters to detect eyelashes in the polar domain and removing them with median filters.

Ma et al. proposed a full framework for iris recognition system as well in [4]. Although it addressed many issues in iris recognition, it does not mention any specific solution for occlusion detection. The only thing related to occlusion is that they discard the lower part of the normalized iris texture and focus only on the more discriminative regions. This scheme is equivalent to automatically assuming the lower part of iris texture is occluded. A similar assumption was proposed in the work of Tisse et al in [5].

Kong and Zhang proposed a model for detecting the eyelashes and specular reflections in [6]. They proposed to use Gabor filters to detect separable eyelashes and use variance of intensity in the local window to locate clusters of multiple eyelashes. After that, the connective criterion is enforced to enhance the robustness of the algorithm. For strong specular reflections, they just used a hard threshold on pixel intensity to identify it. For weak specular reflection, they used mean and standard deviation of the local window as a adaptive threshold for pixel intensity to classify.

In [7], Krichen et al. proposed a probabilistic approach for iris quality measure. They compared the performance of the Gaussian Mixture Model with Fourier-based methods, wavelet-based methods and active contour based methods. The iris masks estimated by their method seem to be local patch-based, not pixel-based. In [8], Thornton proposed to use a discriminative learning method based on FLDA to estimate iris masks in the polar domain.

III. Methodology

In this paper, we focus on the problem of refining the existing iris masks by a learning-based method. Assume we have at least two training images for each class of iris, and we also have roughly estimated iris masks for both of them in polar domain. By maximizing mutual information between every pair of images in training data, we will be able to refine the iris masks of each iris image to a much better quality. In this section, we demonstrate our idea by using only two training images, but the same procedure can be applied to more number of images.

Suppose these two iris images in polar domain are denoted as X and Y, and their corresponding masks are M_X and M_Y. The mutual information between these two masks can be defined as (1)

$$I(M_X; M_Y) = \sum_{y \ni M_Y} \sum_{x \ni M_X} P(x,y) log \frac{P(x,y)}{P(x)P(y)} \quad (1)$$

where x and y represents each pixel on the mask M_X and M_Y in polar domain. Intuitively, the regions on X which have high resemblance to Y are the regions which are more resistant to noises and image distortion. Therefore, we propose to recover the refined mask RM by maximizing the mutual information between X and Y:

$$\begin{aligned} RM &= \arg\max_X I(X; M_Y) \quad (2) \\ &= \arg\max_X \sum_{y \ni M_Y} \sum_{x \ni X} P(x,y) log \frac{P(x,y)}{P(x)P(y)} \end{aligned}$$

Assuming all bits on the polar plane are equally likely to be discriminative regions, therefore, $P(x)$ has a flat distribution across the whole plane. Then (2) can be rewritten as

$$RM = \arg\max_X \sum_{y \ni M_Y} \sum_{x \ni X} P(x,y) log \frac{P(x,y)}{P(y)} \quad (3)$$

By definition of conditional probability, we can rewrite (3) into (4)

$$RM = \arg\max_X \sum_{y \ni M_Y} \sum_{x \ni X} P(x|y)P(y) log P(x|y) \quad (4)$$

We model the probability $P(y)$ by the given roughly estimated mask M_Y, set $P(y) = 1$ if and only if pixel value of y equals one on M_Y and $P(y) = 0$ if and only if pixel value of y equals zero on M_Y. For $P(x|y)$, we model it with the similarity scores between X and Y, given the mask of M_Y. Specifically, we set $P(x|y) = 1$ if pixel at location x on image X matches the same location on image Y, given the roughly estimated mask M_Y, otherwise, set $P(x|y) = 0$. In other words, we model $P(x|y)$ as a binomial distribution.

IV. Experiments

A. Database

We perform our experiments on NIST ICE database, as described in [9]. We used both subsets of ICE, which is ICE1 and ICE2. The details of ICE1 and ICE2 datasets are given in Table I.

Table I
Statistics about ICE1 and ICE2

Dataset	ICE1	ICE2
Number of classes	124	120
Number of Total Images	1425	1528
Minimal Number of Images per class	1	1
Maximal Number of Images per class	31	31

For every iris class in both datasets, we partitioned the images into two mutually exclusive sets, which are training and test set, with equal number of images for both sets. In the case that there exists only one image for the class, it would be included in training data but not test data. We estimated and refined iris masks on training set, and use training set images (partially or totally) as gallery images, and match them against test set (as probe images).

B. Rough Estimation of Iris Masks

We experimented with three different methods for the goal of roughly estimating initial iris masks, which may be inaccurate and error-prone, and then use the proposed method to enhance the performance of the masks. The three methods are described in following sub-sections.

- Rule-based method
  The first method we tried is a simple rule-based method. The rule-based method we used here is similar to the method described in [6]. Basically it detects whether there is strong variance of pixel intensity on a local window and uses it as a feature to perform classification. It can be illustrated in four steps:
  1) Normalize the image so that the energy of pixel intensity sums up to one.
  2) Compute the mean value of the local 5x5 window centered at each pixel. It is the image of mean.
  3) Compute the global mean and standard deviation of all pixel intensity of the image of mean.
  4) For every pixel on the image of mean, if difference between its intensity value and global mean is less than twice the global standard deviation, classify it as iris texture. Otherwise, classify it as an occluded region.
- FLDA-based method
  The second method we tried is a FLDA-based method, as described in [8]. It can be illustrated in following steps:
  1) Extract features from the normalized iris image. The extracted features include:
     a) the mean intensity value in a small neighborhood of the pixel
     b) the standard deviation of the intensity values in the same neighborhood
     c) the percentage of pixels whose intensity is greater than one standard deviation above the mean of the entire iris plane
     d) the shortest Euclidean distance to the centers of the upper and lower eyelids.
  2) Use FLDA method to reduce dimensionality of the features from four to one.
  3) Set a threshold for the feature value to classify a pixel into either class of "authentic iris" or "occluded region".
- Active Contour based method
  As introduced in Section II, Daugman recently proposed a new iris segmentation, based on technique of active contour (Snake). With the new segmentation scheme, together with other improvement in score normalization, eyelashes detection and off-axis eye normalization, he reported a very high accuracy in ICE database. Therefore, we also implement iris mask estimation algorithm by using Snake, in order to re-evaluate it fairly and compare it with other old method. The Snake method we used here is GVF Snake algorithm, as described in [10].

C. Refining Iris Masks

After initial estimation of iris mask is done, we refined the initial iris masks with the proposed method. Since the separation of training data and test data is a random process, to make the experimental results fair and reasonable, we rerun the whole experiments randomly for ten times. This will eliminate any factors that bias the performance due to manipulation of the selection of training data. The experimnetal results reported in Section IV-D is derived by accumulating all Hamming Distance from the ten-fold cross validation for all authentic and imposter comparisons.

For each experimental setting (one specific initial mask estimation algorithm, on one specific database), we ran three sub-experiments and use ROC curve to benchmark their performance. First, we benchmark the performance of the initial mask estimation algorithm, without using proposed method to refine the mask. This setting is called "baseline" in subsequent sections. Second, we refined the masks we got in the baseline method, and randomly pick one training image per class to be the gallery image, and use all untrained images as probe images. This setting is called "Algorithm 1" in subsequent sections. Last, we use every image of every class in training data as gallery images, and use all untrained images as probe images. This setting is called "Algorithm 2" in subsequent sections.

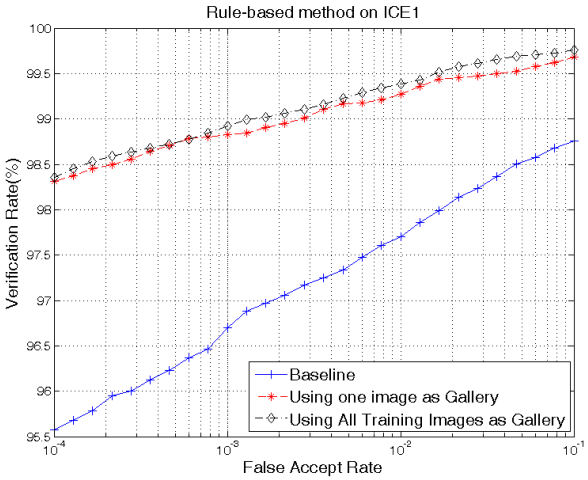
The reason we want to fine-tune and split our proposed method into "Algorithm 1" and "Algorithm 2" is that we would like to push the algorithm to the limit and see how much improvement of matching accuracy we can achieve even in the case when only one training image is selected to be the gallery image. We have a feeling that the proposed algorithm would enhance the matching accuracy to a significant level so that even there is only one training image per class we could still get high performance. We try to validate this assumption through the experiments with "Algorithm 1".

For iris feature extraction, we use Libor Masek's Matlab implementation of Daugman's algorithm, which can be freely downloaded, as stated in [11]. Iris segmentation and normalization is achieved with the method used in [8].

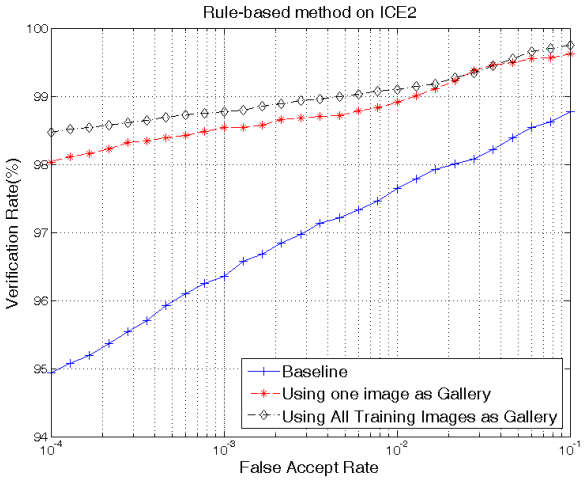
D. Results

As stated in Section IV-C, we present our experimental results in ROC curves with different experimental settings. The results are presented in two parts. First, we plot the ROC curves to illustrate the performance of baseline, Algorithm 1 and Algorithm 2, with respect to each initial iris mask estimation method, on a given database. According to Section IV-B, we have three different methods for estimating initial iris mask, and there are two database (ICE1 and ICE2). Therefore, for this part of result presentation, we have totally six ROC plots. They are shown in Fig. 2.

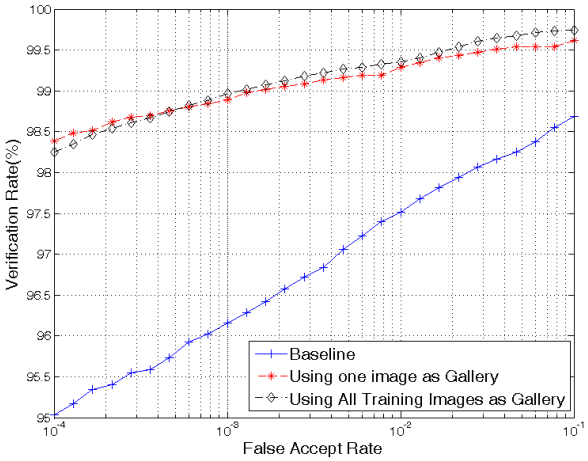
In the second part of our result presentation, we would like to see that before and after applying our proposed algorithm (Algorithm 1 and 2), what is the relative performance of all three initial estimation method? This aspect of comparison would give us an idea whether our proposed algorithm (Algorithm 1 and 2) favor one of the particular initial estimation



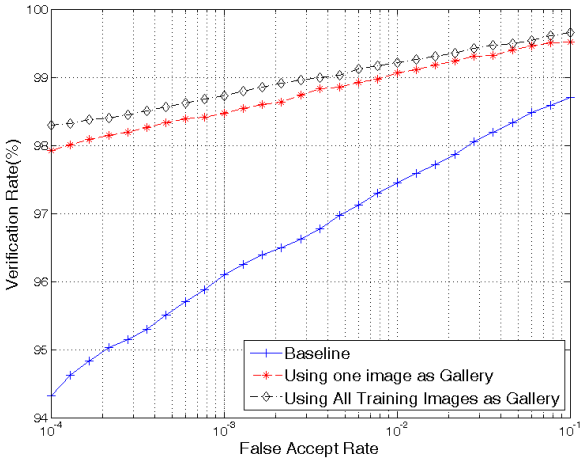
(a)



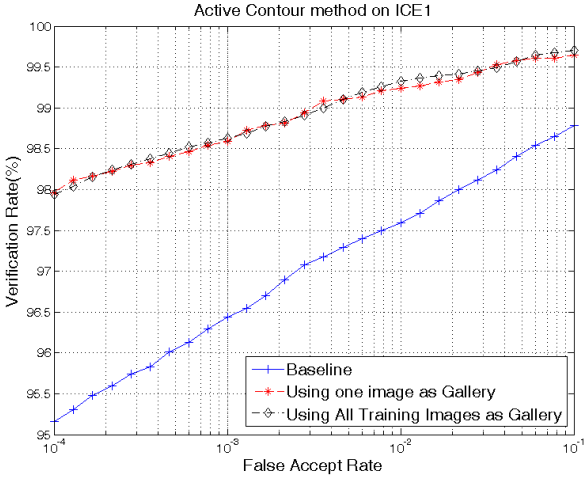
(b)



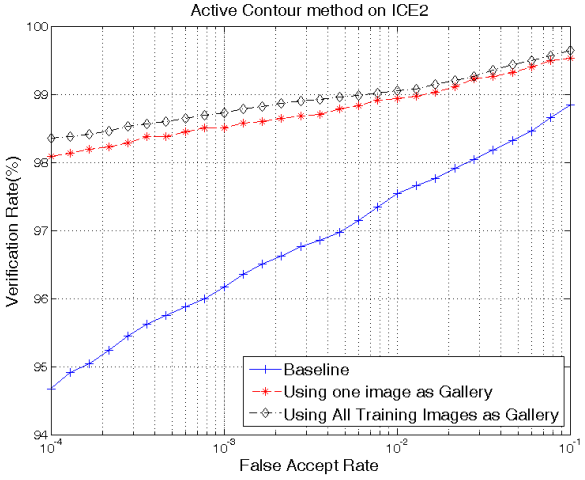
(c)



(d)

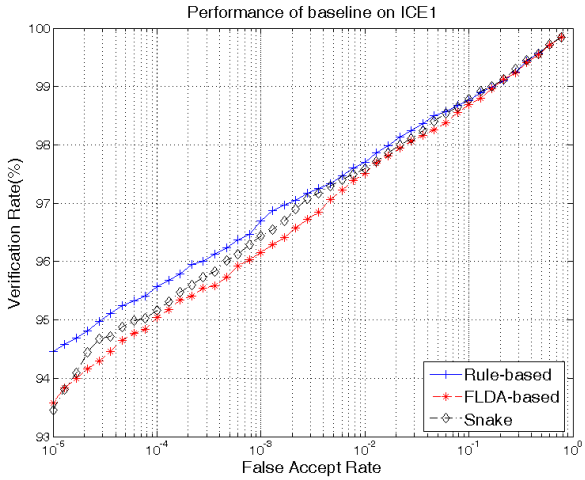


(e)

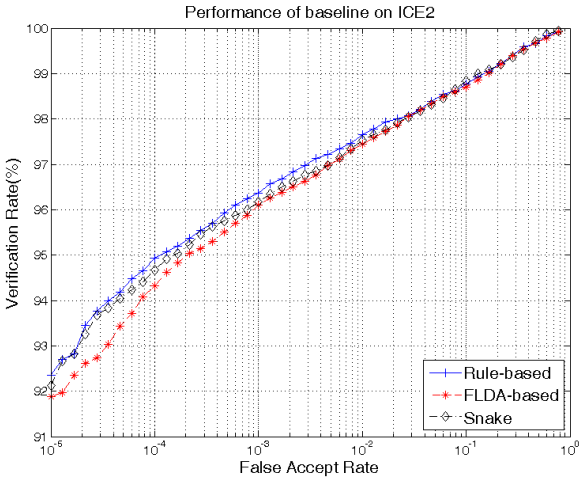


(f)

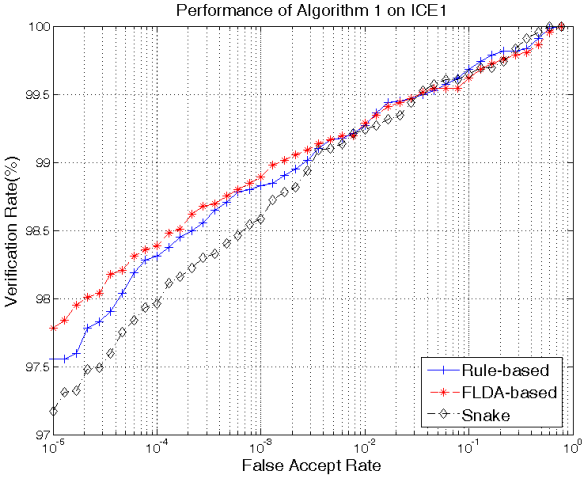
Figure 2. ROC curves for different iris mask estimation algorithm, under different database. (a) Rule-based iris mask estimation algorithm, on ICE1 database. (b) Rule-based iris mask estimation algorithm, on ICE2 database. (c) FLDA-based iris mask estimation algorithm, on ICE1 database. (d) FLDA-based iris mask estimation algorithm, on ICE2 database. (e) Active Contour based iris mask estimation algorithm, on ICE1 database. (f) Active Contour based iris mask estimation algorithm, on ICE2 database.



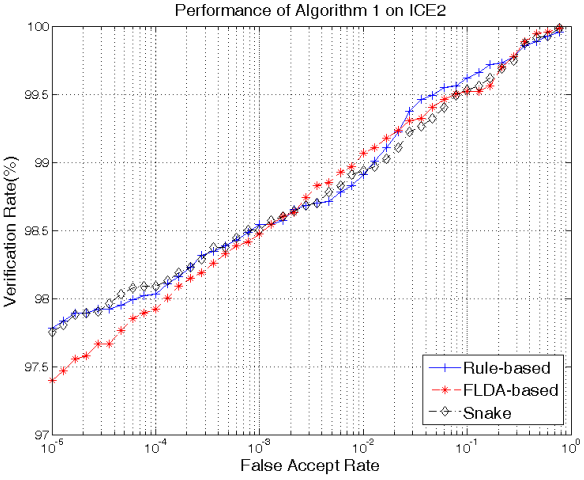
(a)



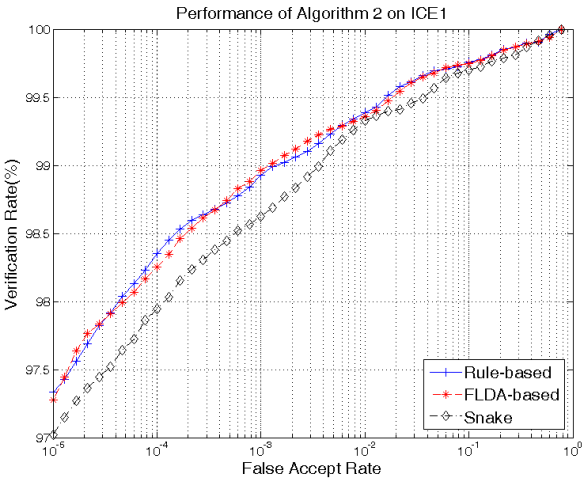
(b)



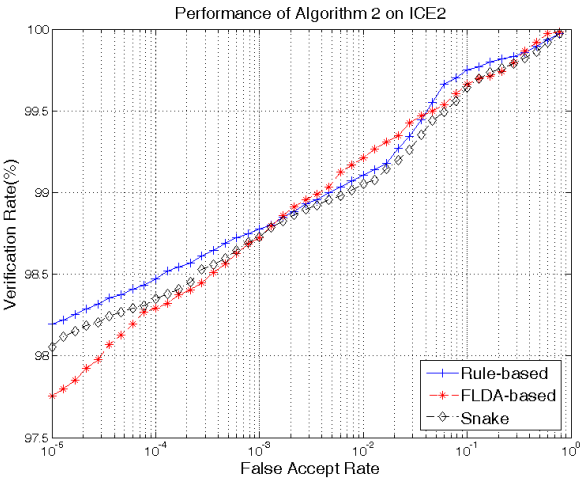
(c)



(d)



(e)



(f)

Figure 3. ROC curves for benchmark the performance of baseline, Algorithm 1 and 2, under different database. Within each plot, performance of three initial iris mask estimation methods is shown. (a) Baseline performance for all three methods, on ICE1 database. (b) Baseline performance for all three methods, on ICE2 database. (c) Performance of all three methods, after improved by Algorithm 1, on ICE1 database. (d) Performance of all three methods, after improved by Algorithm 1, on ICE2 database. (e) Performance of all three methods, after improved by Algorithm 2, on ICE1 database. (f) Performance of all three methods, after improved by Algorithm 2, on ICE2 database.

Table II
NUMERIC PERFORMANCE BENCHMARKS FOR THE THREE DIFFERENT SETTINGS, AVERAGED OVER TEN ITERATIONS, ON ICE1 DATASET

Experimental Settings	Baseline				Algorithm 1				Algorithm 2			
Initial Mask Estimation Method	EER	FRR@.1	FRR@0	FR	EER	FRR@.1	FRR@0	FR	EER	FRR@.1	FRR@0	FR
Rule-based	1.9	3.1	9.3	2.4	0.81	1.1	2.3	4.3	0.7	1.0	3.1	4.4
FLDA-based	2.1	3.7	9.3	2.3	0.8	1.0	2.1	4.2	0.7	1.0	3.7	4.3
Active Contour based	2.0	3.4	10.8	2.4	0.8	1.3	2.9	4.1	0.7	1.3	3.4	4.2

Table III
NUMERIC PERFORMANCE BENCHMARKS FOR THE THREE DIFFERENT SETTINGS, AVERAGED OVER TEN ITERATIONS, ON ICE2 DATASET

Experimental Settings	Baseline				Algorithm 1				Algorithm 2			
Initial Mask Estimation Method	EER	FRR@.1	FRR@0	FR	EER	FRR@.1	FRR@0	FR	EER	FRR@.1	FRR@0	FR
Rule-based	2.0	3.4	9.3	2.3	1.1	1.5	2.3	4.4	0.9	1.2	2.3	4.6
FLDA-based	2.1	3.7	12.9	2.2	0.9	1.5	2.9	4.3	0.8	1.2	2.7	4.5
Active Contour based	2.1	3.6	10.7	2.3	1.1	1.4	2.4	4.3	1.0	1.2	2.3	4.4

method, and we may get more insight about how to choose an initial method wisely. The results are shown in Fig. 3.

For every settings, besides ROC curves, we also measure the numeric performance benchmarks. The benchmarks we measured are: (1) Equal Error Rate (EER); (2) False Reject Rate when False Accept Rate (FAR) equals to 0.1%, denoted as FRR@.1; (3) False Reject Rate when False Accept Rate (FAR) equals to 0%, denoted as FRR@0; (4) Fisher Ratios (FR). The results are shown in Table II and III.

V. DISCUSSION AND CONCLUSIONS

A. Effectiveness of the proposed algorithm

All ROC plots in Fig. 2 consistently confirmed the effectiveness of the proposed algorithm. For every subplot in Fig. 2, one can see that the performance of Algorithm 2 is better than Algorithm 1, which in turn is better than the baseline. This result is perfectly consistent with our assumption, which is that our proposed method will improve the accuracy of the iris mask, which in turn, improves the accuracy of the matching experiment. Also, using more training images as gallery set is definitely better than using less images, which is in accordance with our general knowledge.

B. Algorithm 1 vs. 2

ROC plots in Fig. 2 also show the comparison of the accuracy enhancement through Algorithm 1 and Algorithm2. Although in Fig. 2 (a)(b)(d) and (f), Algorithm 2 is constantly better than Algorithm 1, the lead is very small. Also, in Fig. 2(c) and (e), the perfomance of these two algorithms seem not differ too much. It tells us that even when we randomly pick only one image to be in the gallery set, by using our proposed algorithm to refine the iris mask, one can still achieve high recognition rate. This is a big advantage of the proposed algorithm, especially when we deal with large-scale iris recognition task. If the number of iris classes become larger, by using smaller number of gallery images, one can significantly increase the matching speed, which is an important issue for practical online iris recognition system.

C. Choice of initial iris mask estimation algorithm

From the ROC plots in Fig. 3, we can get insight about how to choose the initial iris mask estimation algorithm. First, Fig. 3(a) and (b) show that among these three initial estimation methods, rule-based method performs the best, and FLDA-based method performs worst. But after applying proposed method, the performance becomes similar among the three. It tells us that no matter which initial estimation algorithm we started with, after applying the proposed refinement algorithm, the iris masks can be improved to approximately the same level of accuracy. This gives us freedom to choose whatever algorithm which is convenient for us. This is also an advantage of the proposed algorithm.

D. Conclusion

The experimental results have shown that our proposed iris mask refinement algorithm is effective and useful. It can refine the iris mask to a much higher quality and improve the iris recognition rate significantly. Moreover, the proposed method is easy to implement, and can co-work with any existing iris mask estimation algorithm. The proposed algorithm can be implemented into the existing iris recognition system as an additional layer between iris mask generation and matching, making it suitable to be used in practical situation without large scale modification of the original system.

One limitation of our proposed method is that it required at least two training images for the same iris class in order to estimate mutual information. This limitation should be easily overcome because modern iris recognition system usually would take more than one iris image during enrollment stage. This fact can be verified on the obervation that most newly released iris database contain more than one image per class.

VI. FUTURE WORKS

Our future work is to test the proposed method on more database, and with more different kinds of iris mask estimation algorithm. We can also explore more on the properties of regions of the refined iris mask to see if there is any interesting characteristics shared among them.

REFERENCES

[1] J. Daugman, "How iris recognition works," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 14, no. 1, pp. 21–30, Jan. 2004.

[2] ——, "New methods in iris recognition," *Systems, Man, and Cybernetics, Part B, IEEE Transactions on*, vol. 37, no. 5, pp. 1167–1175, Oct. 2007.

[3] D. Zhang, D. Monro, and S. Rakshit, "Eyelash removal method for human iris recognition," *Image Processing, 2006 IEEE International Conference on*, pp. 285–288, Oct. 2006.

[4] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Personal identification based on iris texture analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1519–1533, 2003.

[5] C. Tisse, L. Martin, L. Torres, and M. Robert, "Person identification technique using human iris recognition," 2002. [Online]. Available: citeseer.ist.psu.edu/tisse02person.html

[6] W. Kong and D. Zhang, "Accurate iris segmentation based on novel reflection and eyelash detection model," *Intelligent Multimedia, Video and Speech Processing, 2001. Proceedings of 2001 International Symposium on*, pp. 263–266, 2001.

[7] E. Krichen, S. Garcia-Salicetti, and B. Dorizzi, "A new probabilistic iris quality measure for comprehensive noise detection," *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on*, pp. 1–6, Sept. 2007.

[8] J. Thornton, "Matching deformed and occluded iris patterns: a probabilistic model based on discriminative cues," *PhD thesis, Carnegie Mellon University*, 2007.

[9] "Iris challenge evaluation," *National Institute of Standards and Technology, http://iris.nist.gov/ICE/*, 2006.

[10] C. Xu and J. Prince, "Snakes, shapes, and gradient vector flow," *Image Processing, IEEE Transactions on*, vol. 7, no. 3, pp. 359–369, Mar 1998.

[11] L. Masek and P. Kovesi, "Matlab source code for a biometric identification system based on iris patterns," *The School of Computer Science and Software Engineering, The University of Western Australia*, 2003.