Fusion of Face and Iris Biometrics on Mobile Devices Using Near-infrared Images

Qi Zhang[1,2](✉), Haiqing Li[1], Man Zhang[1], Zhaofeng He[1], Zhenan Sun[1], and Tieniu Tan[1,2]

[1] Center for Research on Intelligent Perception and Computing, Institute of Automation, Chinese Academy of Sciences, Beijing, China
[2] College of Engineering and Information Technology, University of Chinese Academy of Sciences, Beijing, China
qi.zhang@nlpr.ia.ac.cn

Abstract. With the wide use of cell phones and tablets, large amounts of private data are stored on mobile devices and personal information security has become a growing concern. Biometrics is able to provide encouraging personal recognition solutions to strengthen the security. This paper proposes a multimodal biometric system for mobile devices by fusing face and iris modalities. Face images are aligned according to eye centers and then represented by histograms of Gabor ordinal measures (GOM). Iris images are cropped from face images and represented by ordinal measures (OMs). Finally, the similarity scores produced by face and iris features are combined in the score level. Experiments are conducted on the CASIA-Mobile database which includes 1400 images of 70 Asians. The proposed system achieves impressive results and demonstrates a promising solution for personal recognition on mobile devices.

Keywords: Face recognition · Iris recognition · Multimodal biometric · Near-infrared · Mobile devices

1 Introduction

Mobile devices such as cell phones and tablets have been widely used for social communication, online shopping and banking. Unprecedented amounts of personal data are stored on mobile devices and facing serious security challenges. It is urgent to build reliable and friendly personal recognition systems for sensitive data protection. While traditional passwords and personal identification numbers (PINs) are easy to crack by guessing or by dictionary attacks, biometrics provides encouraging personal recognition solutions with benefits of its high universality and distinctiveness. Fingerprint recognition is currently available on many mobile devices such as Apple iPhone 6 and Huawei Ascend Mate 7. A number of face recognition applications have been released on line. However, fingerprints left on screens can be easily replicated for spoof attacks and the accuracy of visible light face recognition drops dramatically in complex illumination environment. Fusion of face and iris biometrics using near-infrared (NIR)

J. Yang et al. (Eds.): CCBR 2015, LNCS 9428, pp. 569–578, 2015.
DOI: 10.1007/978-3-319-25417-3_67

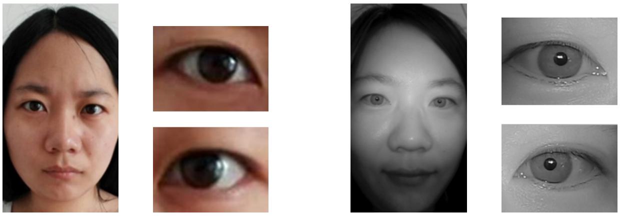
images will overcome these limitations. Firstly, iris imaging does not require contact so that irises are difficult to be replicated. Secondly, NIR illumination is robust to visible light variations and reduces intra-class differences. Finally, face and iris images can be captured simultaneously and possess complementary identity information that will improve the accuracy of a single modality after fusion.

Table 1 summaries recent work on face and iris recognition on mobile devices. Most existing work is based on visible light imagery [1–5]. Marsico et al. [1] implement an embedded biometrics application by fusing face and iris modalities at the score level. After that they build a visible wavelength iris image database named MICHE-I using three different mobile devices under uncontrolled settings [2]. The database collects 3732 images from 92 subjects and has been tested by several algorithms, including spatial histograms [3] and deep sparse filtering [4]. In [4], Raja et al. release a small iris images dataset which consists of 560 images from 28 subjects. Santos et al. [5] build an iris and periocular dataset using four different mobile devices for cross-sensor recognition. All the subjects in the above datasets are Caucasians whose iris texture is clearly displayed under visible light. But as shown in Fig. 1, Asian irises reveal rich texture features only in NIR light due to high density of melanin pigment. Park et al. [6] use a NIR illuminator and a NIR pass filter for iris imaging on mobile phones. 400 face images of 100 subjects are captured in the experiment but not released to the public. In addition, the face information is ignored during recognition.

Table 1. Related work and the corresponding databases.

Contributors	Database	Devices	Illumination	Subjects	Year
BIPLab [2]	MICHE	iPhone5; SamsungGalaxyS4; SamsungGalaxyTab2	visible light	92 Caucasians 3732 images	2015
NISlab [4]	VSSIRIS	iPhone 5S; Nokia Lumia 1020	visible light	28 Caucasians 560 images	2015
Santos et al. [5]	CSIP	Sony Ericsson Xperia Arc S; iPhone 4; ThL W200; Huawei U8510	visible light	50 Caucasians 2004 images	2015
Park et al. [6]	Not released	Samsung SPH-S2300	NIR light	70 Asians 30 Caucasians 400 images	2008

In this paper, we fuse face and iris biometrics on mobile devices for higher security and ease of use. A portable NIR iris imaging module is connected to mobile devices via micro USB to capture rich texture features of Asian irises. The module can capture face and iris images simultaneously and enable us to fuse these two modalities. Face and iris images are first aligned and normalized, and then represented by histograms of Gabor ordinal measures (GOM) [7] and ordinal measures (OMs) [8] respectively. Their similarity scores are added



(a) visible images (b) near-infrared images

Fig. 1. Comparison of visible and near-infrared light iris images. The texture of NIR images is much richer than that of visible images.

together to make the final decision. In order to test our algorithm, the CASIA-Mobile database which consists of 1400 images from 70 Asians is constructed and will be released to the public in the near future. The remainder of this paper is organized as follows. In Section 2, the technical details are described. The experimental results are presented in Section 3. Finally, Section 4 concludes this paper and outlines the future work.

2 Technical Details

2.1 The NIR Iris Imaging Module

The NIR iris imaging module we employed composes of a NIR camera and several NIR illuminators. It is small (about 5cm $\times$ 2cm $\times$ 1cm in Width $\times$ Height $\times$ Thickness) and therefore can be conveniently attached to a mobile phone through micro USB. The module captures valid images at about 25cm standoff distance. The resolution of the whole image is 1080$\times$1920 while the diameter of an iris is about 110 pixels. A face image and two iris images cropped from the face image have been shown in Fig. 1(b).

2.2 Image Preprocessing

In order to reduce the intra-class differences caused by pose variations, face images are cropped and downsampled to 128 $\times$ 128 pixels by transforming eye centers to $(22, 36)$ and $(108, 36)$ with similarity warp. The eye centers of face images are determined by eye detection and iris segmentation. First, the coarse eye regions are detected by Adaboost eye detectors [9]. Then the eye centers and iris boundaries are accurately localized by our previous method [10]. Fig. 2 shows some examples of face alignment. The NIR illuminators are mainly used for iris imaging so that mouths are darker than noses. Only the eye and nose regions are retained to reduce the effects of uniform illumination and barrel distortion. Afterwards, the circular iris is unfolded from polar coordinates into a 70 $\times$ 540 rectangle using the rubber sheet model [11].

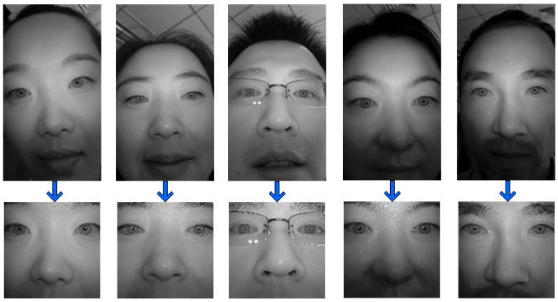


Fig. 2. Examples of face alignment.

2.3 Face Feature Analysis

The method we use for face feature extraction is local feature descriptor, which can extract the detailed information that facilitates the recognition. Two most popular local image descriptors: Gabor filters [12] and OMs are used. The definition of Gabor filter is shown in Eq. (1). The OMs can be expressed as Multi-Lobe Differential Filters (MLDF), which is shown in Eq. (2).

$$G(x,y) = exp(-\frac{1}{2}(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2})) \cdot exp(-2\pi \jmath(\mu_0 x + \nu_0 y)) \tag{1}$$

where (μ_0, ν_0) is the center frequency and σ_x, σ_y is the standard deviation of x and y direction.

$$MLDF = C_m \sum_{i=1}^{N_m} \frac{1}{\sqrt{2\pi}\delta_{mi}} exp[\frac{-(X-\mu_{mi}^2)}{2\delta_{mi}^2}] - C_n \sum_{j=1}^{N_n} \frac{1}{\sqrt{2\pi}\delta_{ni}} exp[\frac{-(X-\mu_{nj}^2)}{2\delta_{nj}^2}] \tag{2}$$

where μ and δ are used to describe the size of each lobe. N_m and N_n is the number of positive and negative lobe, respectively.

The Gabor filter has the advantage in distinctiveness and the ordinal measures have the merit of robustness. Therefore, we integrate these two filters together to achieve a better recognition result. Face images are represented by GOM. Firstly, multi-channel Gabor filters are applied on the normalized face images. Five frequencies and eight orientations Gabor filters are used. Therefore, forty Gabor feature images are generated. Subsequently, four di-lobe and four tri-lobe ordinal measures, whose orientation values are 0°, 45°, 90°, 135°, are used on the phase, magnitude, real and imaginary parts of the Gabor images, respectively. There are totally 5760 dimensional GOM for each feature map block. It is necessary to reduce the feature dimensionality. The widely used linear discriminant analysis (LDA) approach is applied in our paper. LDA is a supervised method aims to find a linear combination of features that best explains the data and can be used for dimensionality reduction. The main idea of LDA is finding a line to project the data that maximizes the between class distance and minimizes the within class variance. The optimal objection function of LDA is shown in Eq. (3).

$$J(\omega) = \frac{\omega^T S_B \omega}{\omega^T S_w \omega} \tag{3}$$

where S_B is between class scatter matrix and S_w is within class scatter matrix. The goal is to maximize the $J(\omega)$.

2.4 Iris Feature Analysis

The robust local image descriptor OMs are used for iris feature extraction, which are widely used in iris recognition [13]. The OMs have many parameters, which differ in scale, orientation, location, lobe numbers, distance, etc. After comprehensive consideration, the di-lobe and tri-lobe ordinal filters are jointly used. Totally, there are 830 various OMs.

In order to get more texture information, the normalized iris image is divided into multiple regions. Usually, the regions adjacent to pupil contain more texture information than other regions. Therefore, various regions of the iris image should be treated differently. In our study, feature extraction is implemented on different regions with various filters. In total, 47,089 regional OMs are extracted to form a huge feature pool.

Taking the time cost into consideration, feature selection is necessary. The number of selected features is usually much smaller than the huge number of candidates, which can speed up the recognition processes. Choosing the distinctive features manually is time-consuming and usually is not the optimal option for a certain database. AdaBoost is used in this paper. Compared with other feature selection methods, AdaBoost has the advantages in high speed, simplicity, classification and feature selection at the same time, etc. AdaBoost algorithm chooses a complementary ensemble of weak learners in a greedy way. It selects a new feature on the reweighted samples. The reweighting strategies are that gain weight of the misclassified examples and lose weight of the correctly classified ones. In this way, future weak learners pay more attention to the examples that are misclassified by previous weak learners. In training process, AdaBoost selects only those features known to improve the predictive power of the model. Thus, it can reduce feature dimensionality. Among the AdaBoost algorithms, GentleBoost outperforms others in that it is more robust to noisy data and less emphasis on the outliers [14]. GentleBoost is used in this paper.

The most distinctive features are obtained on the training dataset using the GentleBoost algorithm. Based on our previous study [15], the recognition performance increases as the number of features increases from zero. However, when the feature number reaches 15, the performance achieves stability. Therefore, in our study, 15 regional features are selected from 47,089 OMs candidates. Subsequently, only the selected 15 regional features are extracted and evaluated on the testing dataset. Thus it can save time and improve the recognition performance.

2.5 Score Level Fusion by the Sum Rule

The cosine similarity is computed to generate matching scores of two face templates (x_1 and x_2). The equation is shown in Eq. (4).

$$cos\theta = \frac{\sum_{k=1}^{n} x_{1k} \cdot x_{2k}}{\sqrt{\sum x_{1k}^2} \cdot \sqrt{\sum x_{2k}^2}} \tag{4}$$

The matching score of two iris codes is calculated by hamming distance (HD). The HD is transformed to similarity at first and then fused with the cosine similarity using the sum rule. Finally, the matching score is normalized to [0,1]. The score level fusion by the sum rule is shown in Eq. (5).

$$FS = (cos\theta + 1 - HD)/2 \tag{5}$$

where FS denotes fusion score. HD is the matching score of two iris codes.

3 Experimental Results

We verify our method on the new built CASIA-Mobile database. At first, face recognition and iris recognition experiments are performed separately. Subsequently, score level fusion of the two modalities is performed.

3.1 CASIA-Mobile Database

In order to promote the face and iris recognition research on mobile phones, we build the CASIA-Mobile database, which will be provided freely for the public. The CASIA-Mobile database includes 1400 face images and 2800 iris images from 70 Asians. There are 40 males and 30 females. The age of the subjects ranges from 22 to 64, among which less than 30 account for 67% proportion. The gender and age distributions can be seen in Fig. 3. For each subject, 20 face images are captured. The left and right irises can be detected and segmented from one face image.

The features of the CASIA-Mobile database are as follows: it is the first public near-infrared mobile database as far as we know. Besides, all the subjects are Asians. Furthermore, the acquisitions of face and iris images are at the same time, which are convenient and beneficial to use more information and perform a bimodal recognition. The challenges of this database include specular reflection of wearing glasses, low resolution images, illumination variation, occlusions of eyelashes and eyelids, etc.

3.2 Face Recognition

The face images of the first 25 subjects are used for training. Each subject has 20 face images. Therefore, there are altogether 500 images for training. 900 face images from the rest 45 subjects are used for testing. There are totally

8550 intra-class matching and 396000 interclass matching. Receiver operating characteristic (ROC) curve and equal error rate (EER) are used for performance measurements. The smaller EER is, the better the performance is. The ROC curve of the face recognition of the CASIA-Mobile database is shown in Fig. 4. The vertical axis is false reject rate (FRR) and the horizontal axis is false accept rate (FAR). The EER is 0.0313.

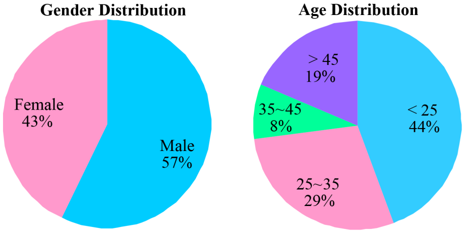


Fig. 3. The gender and age distributions of the subjects in the CASIA-Mobile database.

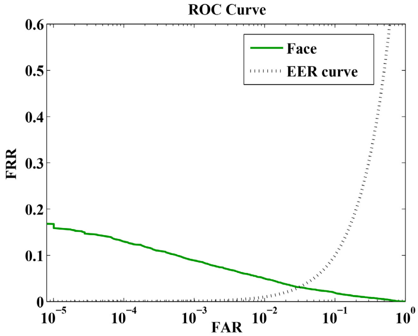


Fig. 4. The ROC curve of the face recognition.

3.3 Iris Recognition

The iris images of the first 25 subjects are used for training. 10 images are chosen for each iris. Therefore, there are altogether 500 images in the training phase. 1800 iris images from the rest 45 subjects (20 images for each iris) are used for testing. In the GentleBoost based feature selection part, there are 2250 intra-class matching scores as positive samples and 4900 inter-class matching scores as negative samples.

The experiments are performed on left and right iris separately. Afterwards, the left and right matching scores are fused by sum rule for better recognition. The EER results are listed in Table 2. The ROC curves of the three experiments are drawn in the same figure, as is shown in Fig. 5.

From the EER results and the ROC curves, it can be concluded that: 1) after score level fusion of the left and right iris, the EER decrease about 50% of the

Table 2. The EER of the iris recognition, where 'Iris LR fusion' represents the score level fusion of left and right iris by the sum rule.

Method	EER
Right iris	0.0757
Left iris	0.0536
Iris LR fusion	0.0389

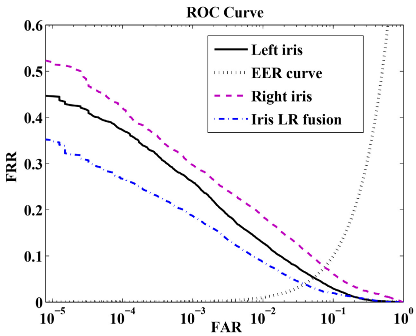


Fig. 5. The ROC curve of the iris recognition, where 'Iris LR fusion' represents the score level fusion of left and right iris by the sum rule.

EER of right iris images; 2) the performance of both the left and right iris are worse than the face recognition result. That is because the iris images are low in resolution. The quality of the captured face images is better than that of the iris images.

3.4 Score Level Fusion

The iris LR fusion experiment is performed at the previous study and achieved better result than the separate iris. In order to further improve the performance, score level fusion of face and iris biometrics experiment is conducted and the sum rule is applied. The EER results are listed in Table 3. The ROC curves of the iris LR fusion, the face and the score level fusion of face and iris biometrics are drawn in the same figure, as is shown in Fig. 6.

Table 3. The EER of the iris and face fusion, where 'Face and iris fusion' represents the score level fusion of face and iris biometrics by the sum rule.

Method	EER
Iris LR fusion	0.0389
Face	0.0313
Face and Iris fusion	0.0076

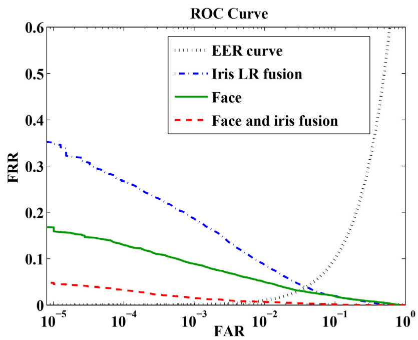


Fig. 6. The ROC curve of iris and face fusion, where 'Face and iris fusion' represents the score level fusion of face and iris biometrics by the sum rule.

From the EER results and the ROC curves, we can draw conclusions that: 1) the EER of the left and right iris fusion is similar to the face recognition result, but the ROC curve is still worse than that of the face recognition; 2) after score level fusion of face and iris biometrics, both the ROC curve and the EER are improved dramatically. It can be inferred that the face and iris modalities have some complementary information and via simple fusion better result can be achieved.

4 Conclusions

In this paper, we have proposed a multimodal biometric system for mobile devices by fusing face and iris modalities. There are three major contributions. First, a portable module has been employed to capture NIR face and iris images simultaneously. Second, an efficient and effective recognition algorithm has been presented. Taking the limited source of mobile devices into consideration, feature dimensionality reduction and feature selection approach are applied to find a small number of features. Face recognition and iris recognition are performed separately and then fused at the score level. Third, a near-infrared database named CASIA-Mobile has been built to promote face and iris recognition algorithm development on mobile devices. After fusion of face and iris modalities, an impressive EER of 0.76% is achieved on the CASIA-Mobile database, which proves the feasibility of our system.

The proposed system can be improved further. The accuracy of iris recognition in our experiments is far from satisfactory due to the low resolution of iris images. The quality of iris images will be improved by better imaging modules and image enhancement algorithms. Apart from the score level fusion, we will also consider other levels of fusion in our future work.

Acknowledgments. This work is supported by the National Natural Science Foundation of China (Grant No.61403389) and the Beijing Nova Programme (Number Z14110-1001814090), China.

References

1. De Marsico, M., Galdi, C., Nappi, M., Riccio, D.: FIRME: face and iris recognition for mobile engagement. Image and Vision Computing **32**(12), 1161–1172 (2014)
2. De Marsico, M., Nappi, M., Riccio, D., Wechsler, H.: Mobile iris challenge evaluation (MICHE)-I, biometric iris dataset and protocols. Pattern Recognition Letters **57**, 17–23 (2015)
3. Barra, S., Casanova, A., Narducci, F., Ricciardi, S.: Ubiquitous iris recognition by means of mobile devices. Pattern Recognition Letters **57**, 66–73 (2015)
4. Raja, K.B., Raghavendra, R., Vemuri, V.K., Busch, C.: Smartphone based visible iris recognition using deep sparse filtering. Pattern Recognition Letters **57**, 33–42 (2015)
5. Santos, G., Grancho, E., Bernardo, M.V., Fiadeiro, P.T.: Fusing iris and periocular information for cross-sensor recognition. Pattern Recognition Letters **57**, 52–59 (2015)
6. Park, K.R., Park, H.A., Kang, B.J., Lee, E.C., Jeong, D.S.: A study on iris localization and recognition on mobile phones. EURASIP J. Adv. Sig. Proc. (2008)
7. Chai, Z., Sun, Z., Mendez-Vazquez, H., He, R., Tan, T.: Gabor ordinal measures for face recognition. IEEE Transactions on Information Forensics and Security **9**(1), 14–26 (2014)
8. Sun, Z., Tan, T.: Ordinal measures for iris recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **31**(12), 2211–2226 (2009)
9. Viola, P., Jones, M.J.: Robust real-time face detection. International Journal of Computer Vision **57**(2), 137–154 (2004)
10. He, Z., Tan, T., Sun, Z., Qiu, X.: Toward accurate and fast iris segmentation for iris biometrics. IEEE Transactions on Pattern Analysis and Machine Intelligence **31**(9), 1670–1684 (2009)
11. Daugman, J.G.: How iris recognition works. IEEE Transactions on Circuits and Systems for Video Technology **14**(1), 21–30 (2004)
12. 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), 1148–1161 (1993)
13. Wang, L., Sun, Z., Tan, T.: Robust regularized feature selection for iris recognition via linear programming. In: International Conference on Pattern Recognition, pp. 3358–3361 (2012)
14. Friedman, J., Hastie, T., Tibshirani, R.: Additive logistic regression: a statistical view of boosting. The Annals of Statistics **28**(2), 337–407 (2000)
15. Sun, Z., Wang, L., Tan, T.: Ordinal feature selection for iris and palmprint recognition. IEEE Transactions on Image Processing **23**(9), 3922–3934 (2014)