Palmprint Recognition Under Unconstrained Scenes

Yufei Han, Zhenan Sun, Fei Wang, and Tieniu Tan

Center for Biometrics and Security Research
National Laboratory of Pattern Recognition, Institute of Automation
Chinese Acdamey of Sciences
P.O.Box 2728, Beijing, P.R. China, 100080
{yfhan,znsun,fwang,tnt}@nlpr.ia.ac.cn

Abstract. This paper presents a novel real-time palmprint recognition system for cooperative user applications. This system is the first one achieving non-contact capturing and recognizing palmprint images under unconstrained scenes. Its novelties can be described in two aspects. The first is a novel design of image capturing device. The hardware can reduce influences of background objects and segment out hand regions efficiently. The second is a process of automatic hand detection and fast palmprint alignment, which aims to obtain normalized palmprint images for subsequent feature extraction. The palmprint recognition algorithm used in the system is based on accurate ordinal palmprint representation. By integrating power of the novel imaging device, the palmprint preprocessing approach and the palmprint recognition engine, the proposed system provides a friendly user interface and achieves a good performance under unconstrained scenes simultaneously.

1 Introduction

Biometrics technology identifies different people by their physiological and behavioral differences. Compared with traditional security authentication approaches, such as key or password, biometrics is more accurate, dependable and difficult to be stolen or faked. In the family of biometrics, palmprint is a novel but promising member. Large region of palm supplies plenty of line patterns which can be easily captured in a low resolution palmprint image. Based on those line patterns, palmprint recognition can achieve a high accuracy of identity authentication.

In previous work, there are several successful recognition systems proposed for practical use of palmprint based identity check [1][2][3], and the best-known is developed by Zhang et al [1]. During image capturing, users are required to place hands on the plate with pegs controlling displacement of hands. High quality palmprint images are then captured by a CCD camera fixed in a semi-closed environment with uniform light condition. To alignment captured palmprint images, a preprocessing algorithm [2] is adopted to correct rotation of those images and crop square ROI (regions of interests) with the same size. Detail about this system can be found in [2]. Besides, Connie et al proposed a peg-free palmprint recognition system [3], which captures palmprint images by an optical scanner. Subjects are allowed to place their hand more freely on the platform of the scanner without pegs. As a result,

Y. Yagi et al. (Eds.): ACCV 2007, Part II, LNCS 4844, pp. 1–11, 2007.

palmprint images with different sizes, translations and rotation angles are obtained. Similar as in [2], an alignment process is involved to obtain normalized ROI images. However, efficient as they are, there are still some limitations. Firstly, some users may feel uncomfortable with pegs to restrict hands during capturing images. Secondly, even without pegs, subjects' hands are required to contact plates of devices or platforms of scanners, which is not hygienic enough. Thirdly, semi-closed image capturing devices usually increase volume of recognition systems, which makes them not convenient for portable use. Thus, it's necessary to improve design of the HCI(human-computer interface), in order to make the whole system easy-to-use.

Recently, active near infrared imagery (NIR) technology has received more and more attention in face detection and recognition, as seen in [4]. Given a near infrared light source shining objects in front of cameras, intensity of reflected NIR light is attenuated at a large scale with distance between objects and the light source increasing. This property provides us a promising solution to eliminate affection of backgrounds when palmprint images are captured under unconstrained scenes. Based on the technology, in this paper, we propose a novel real-time palmprint recognition system. It's designed to localize and obtain normalized palmprint images under clutter scenes conveniently. The main contributions are as followings: First, we present a novel design of portable image capturing device, which mainly consists of two parallel placed web cameras. One is used for active near infrared imagery to localize hand regions. The other one captures corresponding palmprint images in visible light, preparing for further feature extraction. Second, we present a novel palmprint preprocessing algorithm, utilizing color and shape information of hands for fast and effective hand region detection, rotation correction and localization of central palm region. So far as we know, there is no similar work reported in previous literatures.

The rest of paper is organized as follows: Section 2 presents a description of the whole architecture of the recognition system. In Section 3, the design of human computer interface of the system is described in detail. Section 4 introduces ordinal palmprint representation briefly. Section 5 evaluates the performance of the system. Finally, in Section 6, we conclude the whole paper.

2 System Overview

We adopt a common PC with Intel Pentium4 3.0Ghz and 1G RAM as the computation platform. Based on it, the recognition system is implemented using Microsoft Visual C++ 6.0. It consists of five main modules, as shown in Fig.1. After starting the system, users are required to open their hands in a natural manner and place palm regions toward the imaging device at a certain distance between 35 cm and 50 cm from cameras. Surfaces of palms are approximately orthogonal to the optical axis of cameras. In-plane rotation of hands is restricted between -15 degree to 15 degree deviated from vertical orientation. The imaging device then captures two images for each hand by two cameras placed in parallel respectively. One is a NIR hand image with active NIR lighting, the other is a color hand image with background objects, obtained with normal environment lighting condition. Both of them contain complete hand region, see in Fig.2. After that, an efficient palmprint preprocessing

algorithm is performed on the two captured images to obtain one normalized palmprint image quickly, which makes use of both shape and skin color information of hands. Finally, robust palmprint feature templates are extracted from the normalized image using the ordinal code based approach [5]. Fast hamming distance calculation is applied to measure dissimilarity between two feature templates. An example of the whole recognition process could be seen in the supplementary video of this paper.

3 Smart Human-Computer Interface

HCI of the system mainly consists of two parts, image capturing hardware and palmprint preprocessing procedure, as shown in Fig.1. Considering a hand image captured under an unconstrained scene, unlike those captured by devices in [1][2][3], there are not only a hand region containing palmprint patterns, but also background objects of different shapes, colors and positions, as denoted in Fig.2. Even within the hand, there still exits rotation, scale variation and translation of palmprint patterns due to different hand displacements. Thus, before further palmprint feature encoding, HCI should localize the candidate hand region and extract a normalized ROI (region of interest), which contains palmprint features without much geometric deformations.

3.1 Image Capturing Device

Before palmprint alignment, it is necessary to segment hand regions from unconstrained scenes. This problem could be solved by background modeling and subtraction or labeling skin color region. However, both methods suffer from unconstrained backgrounds or varying light conditions. Our design of imaging device aims to solve the problem in a sensor level, in order to localize foreground hand regions more robustly by simple image binarization.

The appearance of the image capturing device is shown in Fig.2(a). This device has two common CMOS web cameras placed in parallel. We mount near infrared (NIR) light-emitting diodes on the device evenly distributed around one camera, similar as in [4], so as to provide straight and uniform NIR lighting. Near infrared light emit by those LEDs have a wavelength of 850 nm. In a further step, we make use of a band pass optical filter fixed on the camera lens to cut off lights with all the other wavelengths except 850nm. Most of environment lights are cut off because their wavelengths are less than 700nm. Thus, lights received by the camera only consist of reflected NIR LED lights and NIR components in environment lights, such as lamp light and sunlight, which are much weaken than the NIR LED lights. Notably, intensities of reflected NIR LED lights are in the inverse proportion to high-order terms of the distance between object and the camera. Therefore, assuming a hand is the nearest one among all objects in front of the camera during image capturing, intensities of the hand region in the corresponding NIR image should be much larger than backgrounds. As a result, we can further segment out the hand region and eliminate background by fast image binarization, as denoted in Fig.2(b). The other

camera in the device captures color scene images, obtaining clear palmprint patterns and reserving color information of hands. An optical filter is fixed on the lens of this camera to filter out infrared components in the reflected lights, which is applied widely in digital camera to avoid red-eye. The two cameras work simultaneously. In our device, resolution of both cameras is 640*480. Fig.2(b) lists a pair of example images, captured by the two cameras at the same time. The upper one is the color image. The bottom one is the NIR image. The segmentation result is shown in the upper row of Fig.2(c). In order to focus on hand regions with a proper scale in further processing, we adopt a scale selection on binary segmentation results to choose candidate foreground regions. The criterion of selection grounds on a fact that area of a hand region in a NIR image is larger if the hand is nearer to the camera. We label all connected binary foreground after segmentation and calculate area of each connected component, then choose those labeled regions with their areas varying in a predefined narrow range as the candidate foreground regions, like the white region shown in the image at the bottom of Fig.2(c).

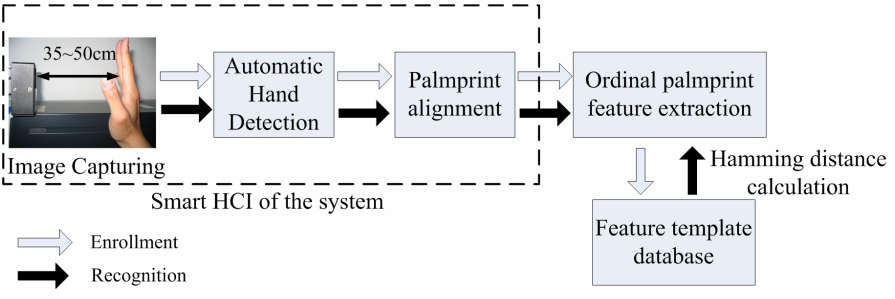

Fig. 1. Flowcharts of the system

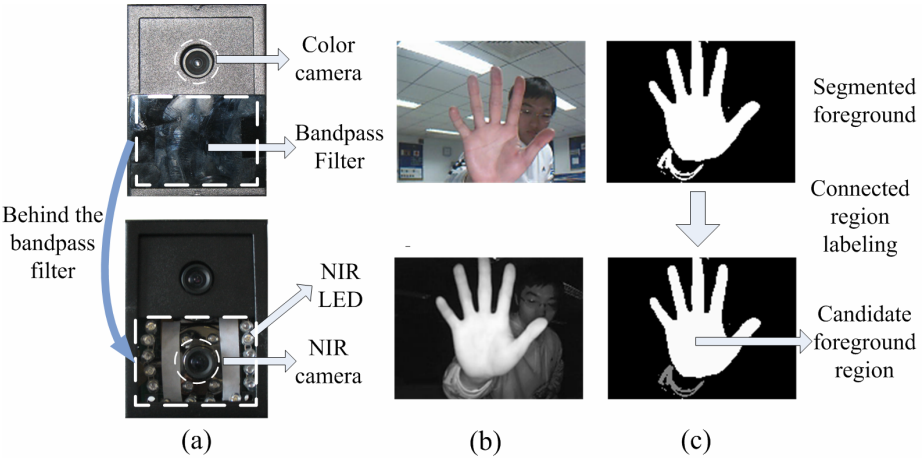

Fig. 2. (a) Image capturing device (b) Pair-wise color and NIR image (c) Segmented fore ground and candidate foreground region selection

3.2 Automated Hand Detection

Hand detection is posed as two-class problem of classifying the input shape pattern into hand-like and non-hand class. In our system, a cascade classifier is trained to detect hand regions in binary foregrounds, based on works reported in [6]. In [6], Eng-Jon Ong et al makes use of such classifier to classify different hand gestures. In our application, the cascade classifier should be competent for two tasks. Firstly, it should differentiate shape of open hand from all the other kinds of shapes. Secondly, it should reject open hands with in-plane rotation angle deviating out of the restricted range. To achieve this goals, we construct a positive dataset containing binary open left hands at first, such as illustrated in Fig.3(a). In order to make the classifier tolerate certain in-plane rotation, the dataset consists of left hands with seven discrete rotation angles, sampled every 5 degree from -15 degree to 15 degree deviated from vertical orientation, a part of those binary hands are collected from [11]. For each angle, there are about 800 hand images with slight postures of fingers, also shown in Fig.3(a). Before training, all positive data are normalized into 50*35 images. The negative dataset contains two parts. One consists of binary images containing non-hand objects, such as human head, turtles and cars, partly from [10]. The other contains left hands with rotation angle out of the restricted range and right hands with a variety of displacements. There are totally more than 60,000 negative images. Fig.3(b) shows example negative images. Based on those training data, we use Float AdaBoost algorithm to select most efficient Haar features to construct the cascade classifier, same as in [6]. Fig.3(c) shows the most six efficient Haar features obtained after training. We see that they represent discriminative shape features of left open hand. During detection, rather than exhaustive search across all positions and scales in [6], we perform the classifier directly around the candidate binary foreground regions

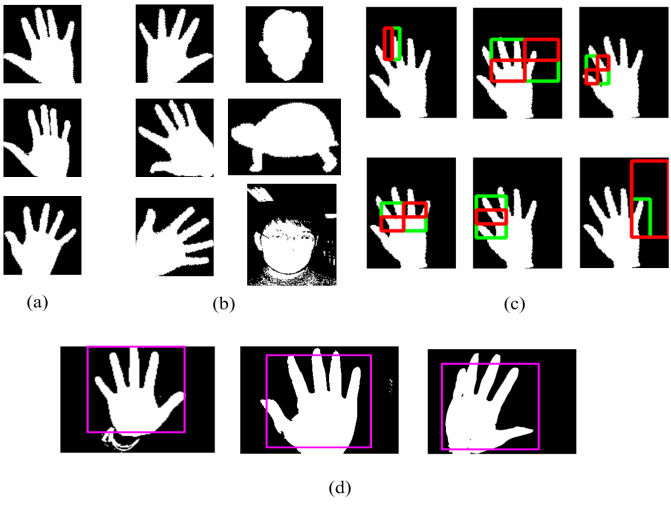

Fig. 3. (a) Positive training data (b) Negative training data (c) Learned efficient Haar features (d) Detected hand region

to search for open left hands with a certain scale. Therefore, we can detect different hands with a relative stable scale, which reduces influence of scale variations on palmprint patterns. Considering mirror symmetry between left and right hands, to detect right hands, we just perform symmetry transform on the images and apply the classifier by the same way on the flipped images. Fig.3(d) shows results of detection. Obtaining detected hand, all the other non-hand connected regions are removed from binary hand images. The whole detection can be finished within 20 ms.

3.3 Palmprint Image Alignment

Palmprint alignment procedure eliminates rotation and translation of palmprint patterns, in order to obtain normalized ROI. Most alignment algorithms calculate rotation angles of hands by localizing key contour points in gaps between fingers [2][3]. However, in our application, different finger displacements may change local contours and make it difficult detect gap regions, as denoted in Fig.4. To solve this problem, we adopt a fast rotation angle estimation based on moments of hand shape.

Given R is the detected hand region in a binary foreground image. Its orientation θ can be estimated by calculating its moments [7]:

$$\theta = \frac{1}{2}\arctan\left(\frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}}\right) \tag{1}$$

$\mu_{p,q}$ (p,q=0,1….) is (p,q) order central moments, which is represented as :

$$\mu_{p,q} = \sum_x \sum_y \left(x - \frac{1}{N}\sum_x \sum_y x\right)^p \left(y - \frac{1}{N}\sum_x \sum_y y\right)^q, (x, y) \in \mathrm{R} \tag{2}$$

Compared with key point detection, moments are calculated based on the whole hand region rather than only contour points. Thus, it is more robust to local changes in contours. To reduce computation cost, the original binary image is down-sampled to a 160*120 one. Those moments are then calculated on the down-sampled version. After obtaining rotation angle θ, the hand region is rotated by $-\theta$ degree to get vertical oriented hands, see in Fig.4. Simultaneously, the corresponding color image is also rotated by $-\theta$, in order to make sure consistency of hand orientations in both two images.

In a further step, we locate central palm region in a vertical oriented open hand by analyzing difference of connectivity between the palm region and the finger region. Although shape and size of hands vary a lot, a palm region of each hand should be like a rectangle. Compared with it, stretched fingers don't form a connective region as palm. Based on this property, we employ an erosion operation on the binary hand image to remove finger regions. The basic idea behind this operation is run length code of binary image. We perform a raster scanning on each row to calculate the maximum length W of connective sequences in the row. Any row with its W less than threshold K_1 should be eroded. After all rows are scanned, a same operation is performed on each column. As a result, columns with their maximum length W less than K_2 are removed. Finally, a rectangular palm region is cropped from the hand. Coordinates (x_p,y_p) of its central point is derived as localization result. In order to

cope with varying sizes of different hands, we choose values of K_1 and K_2 adaptively. Before row erosion, distance between each point in the hand region and nearest edge point is calculated by a fast distance transform. The central point of hand is defined as the one with the largest distance value. Assuming A is the maximum length of connective sequences in the row passing through the central point, K_1 is defined as follows:

$$K_1 = A * p\% \tag{3}$$

p is a pre-defined threshold. $K2$ is defined in the same way:

$$K_2 = B * q\% \tag{4}$$

B is the maximum length of connective sequences in the column passing through the central point after row erosion. q is another pre-defined threshold. Compared with fixed value, adaptive K_1 and K_2 lead to more accurate location of central palm regions, as denoted in Fig.5(b). Fig.5(a) denotes the whole procedure of erosion.

Due to visual disparity between two cameras in the imaging device, we can not use (x_p,y_p) to localize ROI in corresponding color images directly. Although visual disparity can be estimated by a process of 3D scene reconstruction, this approach may lead to much computation burden on the system. Instead, we apply a fast correspondence estimation based on template matching. Assuming C is a color hand image after rotation correction, we convert C into a binary image M by setting all pixels with skin color to 1, based on the probability distribution model of skin color in RGB space [8]. Given the binary version of the corresponding NIR image, with a hand region S locating at (x_n,y_n), a template matching is conducted as in Eq.5, also as denoted in Fig.6:

$$f(m,n) = \sum_x \sum_y [M(x+m, y+n) \oplus S(x, y)], (x, y) \in S \tag{5}$$

$\oplus$ is bitwise AND operator. $f(m,n)$ is a matching energy function. (m,n) is a candidate position of the template. The optimal displacement (x_o,y_o) of hand shape S in M is defined as the candidate position where the matching energy achieves its maximum. The central point (x_c,y_c) of palm region in C can be estimated by following equations:

$$\begin{aligned} x_c &= x_p + x_o - x_n \\ y_c &= y_p + y_o - y_n \end{aligned} \tag{6}$$

Fig. 4. Rotation correction

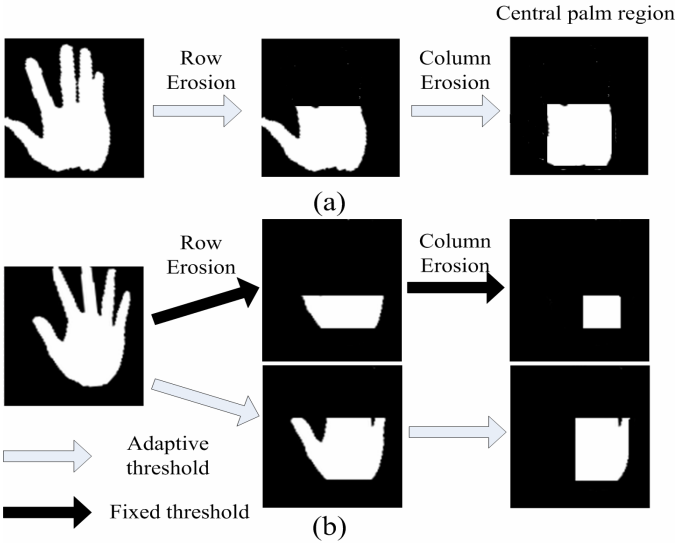

Fig. 5. (a) Erosion procedure (b) Erosion with fixed and adaptive thresholds

With (x_c,y_c) as its center, one 128*128 sub-image is cropped from C as ROI, which is then converted to gray scale image for feature extraction.

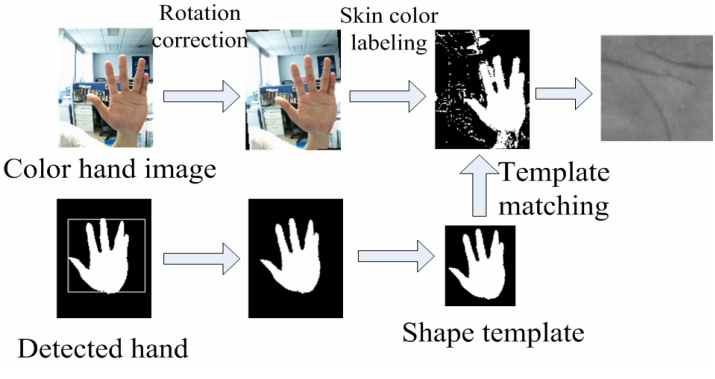

Fig. 6. Translation estimation

4 Ordinal Palmprint Representation

In previous work, the orthogonal line ordinal feature (OLOF) [5] provides a compact and accurate representation of negative line features in palmprints. The orthogonal line ordinal filter [5] $F(x,y,\theta)$ is designed as follows:

$$F(x, y, \theta) = G(x, y, \theta) - G(x, y, \theta + \pi / 2) \tag{7}$$

$$G(x, y, \theta) = \exp[-(\frac{x\cos\theta + y\sin\theta}{\delta_x})^2 - (\frac{-x\sin\theta + y\cos\theta}{\delta_y})^2] \quad (8)$$

$G(x,y,\theta)$ is a 2D anisotropic Gaussian filter, and θ is the orientation of the Gaussian filter. The ratio between δ_x and δ_y is set to be larger than 3, in order to obtain a weighted average of a line-like region. In each local region in a palmprint image, three such ordinal filters, with orientations of 0, $\pi/6$, $\pi/3$ are used in convolution process on the region. The filtering result is then encoded into 1 or 0 according to whether its sign is positive or negative. Thousands of ordinal codes are concatenated into a feature template. Similarity between two feature templates is measured by a normalized hamming distance, which ranges between 0 and 1. Further details can be found in [5].

5 System Evaluation

Performance of the system is evaluated in terms of verification rate [9], which is obtained through one-to-one image matching. We collect 1200 normalized palmprint ROI images from 60 subjects using the system, with 10 images for each hand. Fig.**7** illustrates six examples of ROI images. During the test, there are totally 5,400 intra-class comparisons and 714,000 inter-class comparisons. Although recognition accuracy of the system lies on effectiveness of both alignment procedure of HCI and the palmprint recognition engine, the latter is not the focus of this paper. Thus we don't involve performance comparisons between the ordinal code and other state-of-the-art approaches. Fig.8 denotes distributions of genuine and imposter. Fig.9 shows corresponding ROC curve. The equal error rate [9] of the verification test is 0.54%. From experimental results, we can see that ROI regions obtained by the system are suitable for palmprint feature extraction and recognition. Besides, we also record time cost for obtaining one normalized palmprint image using the system. It includes time for image capturing, hand detection and palmprint alignment. The average time cost is 1.2 seconds. Thus, our system can be competent for point-of-sale identity check.

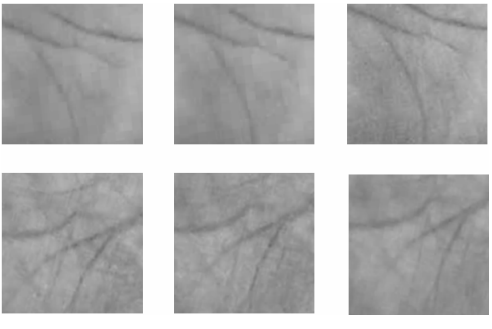

Fig. 7. Six examples of ROI images

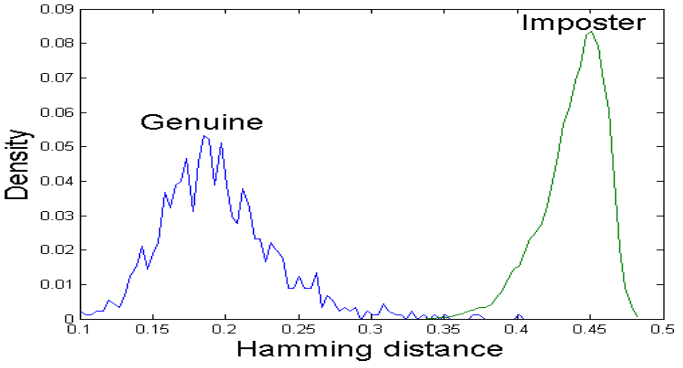

Fig. 8. Distributions of genuine and imposter

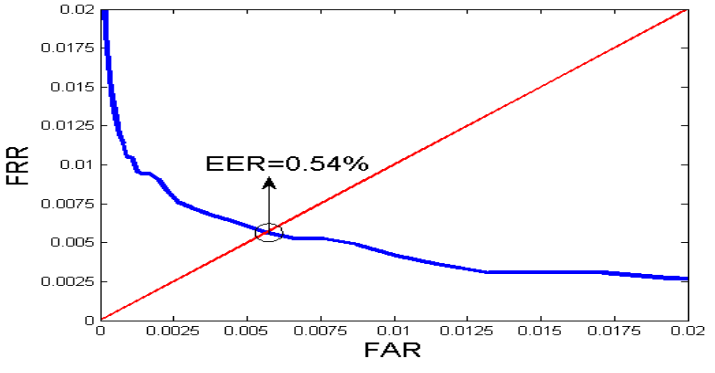

Fig. 9. ROC curve of the verification test

6 Conclusion

In this paper, we have proposed a novel palmprint recognition system for cooperative user applications, which achieves a real-time non-contact palmprint image capturing and recognition directly under unconstrained scenes. Through design of the system, we aim to provide more convenient human-computer interface and reduce restriction on users during palmprint based identity check. The core of HCI in the system consists of a binocular image device and a novel palmprint preprocessing algorithm. The former delivers a fast hand region segmentation based on NIR imaging technology. The latter extracts normalized ROI from hand regions efficiently based on shape and color information of human hands. Benefiting further from the powerful recognition engine, the proposed system achieves accurate recognition and convenient use at the same time. As far as we know, this is the first attempt to solve the problem of obtaining normalized palmprint images directly from clutter backgrounds.

However, accurate palmprint alignment has not been well addressed in the proposed system. In our future work, it's an important issue to improve the performance of the system by reducing alignment error in a further step. In addition,

we should improve the imaging device to deal with influence of NIR component in environment light, which varies much in practical use.

Acknowledgments. This work is funded by research grants from the National Basic Research Program (Grant No.2004CB318110), the Natural Science Foundation of China (Grant No.60335010, 60121302, 60275003, 60332010, 69825105,60605008) and the Chinese Academy of Sciences.

References

1. Zhang, D., Kong, W.K., You, J., Wong, M.: Online Palmprint Identification. IEEE Trans on PAMI 25(9), 1041–1050 (2003)
2. Kong, W.K.: Using Texture Analysis in Biometric Technology for Personal Identification, MPhil Thesis, http://pami.uwaterloo.ca/ cswkkong/Sub_Page/Publications.htm
3. Connie, T., Jin, A.T.B., Ong, M.G.K., Ling, D.N.C.: Automated palmprint recognition system. Image and Vision Computing 23, 501–515 (2005)
4. li, S.Z., Chu, R.F., Liao, S.C., Zhang, L.: Illumination invariant Face Recognition using Near- Infrared Images. IEEE Trans on PAMI 29(4), 627–639 (2007)
5. Sun, Z.N., Tan, T.N., Wang, Y.H., Li, S.Z.: Ordinal Palmprint Representation for Personal Identification. Proc. of IEEE CVPR 2005 1, 279–284 (2005)
6. Ong, E., Bowden, R.: A Boosted Classifier Tree for Hand Shape Detection. In: Proc. of International Conference on Automatic Face and Gesture Recognition, pp. 889–894 (2004)
7. Jain, A.K.: Fundamentals of Digital Image Processing, vol. 07458, p. 392. Prentice Hall, Upper Saddle River, NJ
8. Jones, M.J., Rehg, J.M.: Statistical Color Models with Application to Skin Color Detection. International Journal of Computer Vision 46(1), 81–96 (2002)
9. Daugman, J., Williams, G.: A Proposed Standard for Biometric Decidability. In: Proc. of CardTech/SecureTech Conference, Atlanta, GA, pp. 223–234 (1996)
10. http://www.cis.temple.edu/ latecki/TestData mpeg7shapeB.tar.gz
11. UST Hand Image database, http://visgraph.cs.ust.hk/Biometrics/Visgraph_web/ index.html