Real-time ROI Acquisition for Unsupervised and Touch-less Palmprint

Yi Feng, Jingwen Li, Lei Huang, and Changping Liu

Abstract—In this paper we proposed a novel method to acquire the ROI (Region of interest) of unsupervised and touch-less palmprint captured from a web camera in real-time. We use Viola-Jones approach and skin model to get the target area in real time. Then an innovative course-to-fine approach to detect the key points on the hand is described. A new algorithm is used to find the candidate key points coarsely and quickly. In finely stage, we verify the hand key points with the shape context descriptor. To make the user much comfortable, it can process the hand image with different poses, even the hand is closed. Experiments show promising result by using the proposed method in various conditions.

Keywords—Palmprint recoginition, hand detection, touch-less palmprint, ROI localization.

I. INTRODUCTION

PALMPRINT recognition is a reliable personal identity method. The palm consists of many features, such as principal lines, ridges wrinkles, delta points. These features can be used sufficiently in verification processes. Compared with the other biometric technology, palm print authentication has several advantages. Firstly, the print patterns are unique, and the palm print characteristics are more abundant than fingerprint and iris. Secondly, the iris patterns required high resolution images. While the palmprint recognition use the principal and wrinkles which are also discriminating in low-resolution images. So the capture devise is less expensive than iris recognition. Moreover, palm print recognition system presents much higher user acceptability than iris and fingerprint.

Most current palm print recognition systems have a complex device for controlling the light, background, and hand position [1, 2]. These systems are much larger than fingerprint recognition system. The users must put their hand in the semi-enclosed box, on the sensor or a plate with pegs. It makes user very uncomfortable during identification and causes sanitary issue in the public areas. These reasons greatly limit the application of the palm print recognition system. So it is very important to design a novel system to process the touch-less and unsupervised hand image.

In general, the related research of touch-less and

Yi Feng is with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, CO 100080 China (e-mail: yi.feng@ia.ac.cn).
JingWen Li is with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, CO 100080 China (e-mail: jingwen.li@ia.ac.cn).
Lei Huang is with the Institute of Automation, Chinese Academy of Sciences, CO 100080 China (e-mail: lei.huang@ia.ac.cn).
Changping Liu is with the Institute of Automation, Chinese Academy of Sciences, CO 100080 China (e-mail: changping.liu@ia.ac.cn).

unsupervised palmprint recognition are is not much. In [3, 4, 5], J. Doublet et al. proposed a system that uses color information and neural network model for hand detection; uses shape information and Active Shape Model(ASM) for key points detection. Because the ASM model is learned from the marked hand image, it can't work well when the hand pose changed too much. In [6], Ivan Fratric et al. described a Real-Time Model-Based Hand Localization. Viola-Jones approach was selected for hand candidate detection. During hand localization phase, a large number of hand models were used for matching the candidate. However the shape and the size between individuals are quite different, the hand models can't fit every hand accurately. In [7], Ong et al. designed a touch-less palmprint verification system, which use Gaussian model for hand segmentation and a novel method to find hand key points. In [8], Michal Choras et al. used skin threshold for hand segmentation. Though these approaches [7, 8] are very concise, they may not work in cluster background. In [9], Yufei Han et al. used two parallel placed web cameras for ROI acquisition. One captures the active near infrared imagery for hand detection ROI localization. The other acquired in visible light, prepare for ROI acquisition. During ROI acquisition, they analyzed difference of connectivity between the foreground (hand area) and the background (non-hand area), which is not stable because of hand pose changing.

In this paper we describe the development of a real-time ROI acquisition for unsupervised and touch-less palmprint system which will be used in palmprint recognition. It will not require the user's hand touch any platform. The users just need to put their hand in front of a sensor in the unconstrained scenes. They can open their hand, close their hand or pose others in natural manner, such as illustrated in Fig.1. This system is also robust enough to work in various lighting and cluster background.

Fig.1 Some hand images in our database

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:6, 2011

II. SYSTEM ORGANIZATION

Fig.2 shows the structure of our proposed system for ROI extraction. We use a web camera (resolution of the camera is 640×480) to capture real-time palm print images. The palmprint region is required to be toward the sensor. The position, orientation, and pose of the hand are freely. The distance between the hand and the sensor is around 8 cm - 30 cm. If the hand is too far away from the camera, the handprint will be too fuzzy to recognize. On the other hand, if the hand is very close to the camera, we can't get the whole palmprint. As we see from Fig.1, the challenges of ROI acquisition is that the scale, the position, the orientation, the pose of a hand and the light, the background of a scene are quite different.

Fig.2 Structure of the proposed system.

We design a coarse-to-fine strategy to find the ROI of a hand. After get the frame from the camera, we resize the image to the resolution of 160×120 for saving the process time. Viola-Jones approach was used to discriminate weather the image include a hand because of its speed and good detection rate [10]. Then we use the Bayesian maximum likelihood classifier to find the skin pixels in the image and get a binary skin area image.

After that, we proposed a novel method to find the key points in the skin area coarsely quickly, and then use the shape context descriptor of the key points to verify whether the point is the key point of a hand. This strategy is fully satisfactory, in terms of both efficiency and accuracy.

At last, we get the location of ROI and acquire the ROI of palmprint for recognition with the key point and original hand image.

III. DETECT TARGET AREA

Owing to the system's requirement of real-time, we must use a kind of quick and effective method to discard the non-hand image and segment the target area. So we chose Viola and Jones [10] for detection and skin model for segmentation.

A. Hand detection

The goal of hand detection process is to discriminate weather the image contain a hand. This process should be as fast as possible, so as to be able to process the entire frame in real time. Viola and Jones [10] originally proposed the cascade of boosted classifiers as a real-time general object detector and applied it to face detection. They showed that the system works quite well at various scales and with different backgrounds under various illumination conditions. We apply Lienhart's methods, as

implemented in the OpenCV, Open Computer Vision Library (2006), to the hand detection problem.

B. Skin Segmentation

In the skin segmentation phrase, we constructed skin and non-skin histogram models using a large training set. Based on the Son Lam Phung's result [11] and our own follow-up study we selected the YCbCr color model. Histograms used in the skin detector have two dimensions: Cb and Cr channels. Each axis of the plane is quantized into 32 bins, so that each histogram will have 32^2 (1024) bins. Although the larger histogram sizes tend to perform better, the differences in classification rates for histogram sizes of 256, 128, 64 and 32 bins per channel were quite small [11]. So we chose the 32 bins per channel for saving the memory storage. In the training process, 162902045 hand skin pixels from the 1500 hand images were labeled manually and placed into the skin histogram. 614400000 non-skin pixels from 2000 images that did not contain skin were placed into non-skin histogram. Both histograms are constructed by simply counting the number of pixels which belong to same bin, and they are normalized by the total number of pixels used to construct the histogram. Then we get the discrete class-conditional likelihood: P(CbCr | skin), P(CbCr | non-skin). We derive a skin pixel classifier through the standard likelihood ratio approach [12]. A particular CbCr value is labeled skin if

$$\frac{P(CbCr \mid skin)}{P(CbCr \mid non-skin)} \geq \theta \tag{1}$$

Where $0 \leq \theta \leq 1$ is a threshold which can be adjusted to trade-off between correct detection and false positives. After classify every pixel, we get a skin area binary image in which the skin pixel is 1 and the non-skin pixel is 0.

Then we find the largest connected area of skin pixels as the target region. Due to the distance between the camera and hand is 5 cm – 30 cm, the area of target region should more than one-sixth of the whole image. So we discard the small target area, then the morphological method is used to remove the noise.

IV. KEY POINT DETECTION COARSELY

In palm print recognition, the key point detection plays a very important role for ROI localization. These stable points, such as the points between two finger roots, see in Fig.3 (a), are used to obtain normalized ROI of palmprint which is steady for changes of hade position, hand scale, and hand pose. In [3, 4, 5, 6, 7, 8, 9], they all get the hand key points from the hand contours and required the hand should open. Their approaches cannot process the other pose of hand, such as in Fig.3 (a). We will describe a new method to process the non-open hand images and find key points, which use both the edge map and skin area contour.

A. Find the whole counter of the candidate finger and hand

As shown in Fig.3 (b), we get the external contour from the

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:6, 2011

skin area easily. However it is very different to find the counters of the figure in a close hand , see in Fig.3(c). We calculate the edge map of the original image with the sobel edge detector, which is illustrated in Fig.3 (d). Then we trace the figure counters in the edge which connect to the external counter and have a high gradient value over the threshold. We also use standard edge linking and thinning methods with morphological operators to extracted contours. At last we get the entire counter of the finger and hand, which is shown in Fig.3 (e).
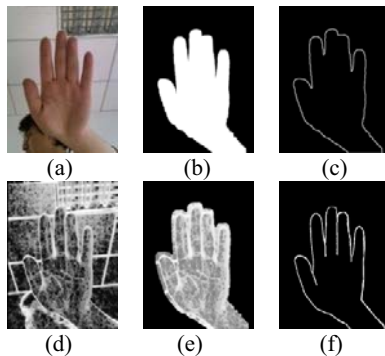
Fig.3 (a) Original image with key points marked red; (b) The target area; (c) The external contour of skin area; (d) The edge map of original image with histogram equalizing; (e) The edge map of the original image in target area with histogram equalizing; (f) The entire counter of the finger and hand.

B. Detect the key point on the counter

Ong et al. proposed competitive hand valley detection (CHVD) algorithm to find detect the key points [7], which cannot process the close hand. Inspired by their thought, we designed a new algorithm.

As shown in Fig.4, there are two kind of key points between the finger roots: finger open and finger together. They are separated into class A and class B. We use Ong's algorithm to find class A points, then trace the entire counter to find candidate class B points. The neighboring points of different distance will be used for discriminating and three checking steps are:

1) Observe the four domain neighbors of the current point. If one of the points upon falls in the finger edge, while the other within the skin area, this point is considered a candidate class B point and process next step, otherwise discard this point.

2) Observe the twelve neighbors ten pixels away from the current point. If there is at least one and no more than three finger edge points upon it, and all of underside consecutive points are in skin area (pixel value = 1), this point is considered a candidate class B point and process next step, otherwise discard this point.

3) To verify the current point, we draw a line perpendicular to the edge. If it passes only consecutive skin area in a certain distance, which mainly depending on the hand scale, the current point will be asserted as a class B point.
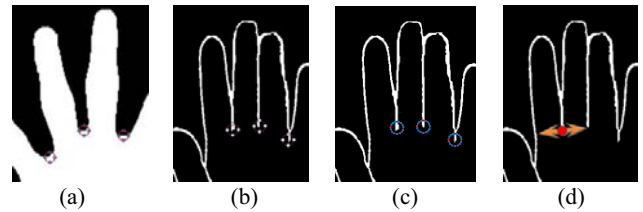
Fig.4 (a) The points in class A; (b) The points in class B and check algorithm of step 1; (c) Check algorithm of step 2; (d) Check algorithm of step 3.

V. VERIFY THE KEY POINTS AND ROI ACQUISITION

Some wrong key points are detected because of the poor quality of the contour. In this phrase, we verify the key points more finely and use the key points to location the ROI of palmprint. In [3, 4, 5, 10], they use deformable templates through matching-based approaches to localization hand accurately. These models cannot process the changes of hand pose. Actually relative position of the hand key points is much stable in various hand positions, which shape context descriptor can represent well.

A. Verify the key points

The shape context descriptor was proposed by Belongie et al. in [12]. It allows to measure shapes similarity by recovering point correspondences between two objects. It has the following invariance properties: translation, because it is based on relative point location; scale, owning to the normalizing the radial distance by the mean (or median) distance between all point pairs; shape, it is robust toward slight shape variation. However, it is not powerful enough to yield reliable point correspondence in cluster scenes and it cost a lot of time for matching the two point sets.

In our system, shape context descriptor is used to verify the hand key points and discard some wrong key points due to the error of skin detection which need not to match all the points on the counter. For each key point on the image, the local edge feature distribution is measured based on the shape context [12]. In our experimental setup, we have chosen 5 bins for log r and 12 bins for θ and normalizing the all radial distance by the mean distance between all the key point pairs.

In the learning stage, we manually marked 1000 key points on 250 hand counter which acquired from our hand image database. Actually there are some differences between the points in different hand position, so we'd better divide them into several sets by their descriptor. After the shape context of each key point was calculated, we cluster these points by means of the k-means algorithm and get the k cluster centers. As shape contexts are distributions represented as histograms, it is natural to use the test statistic to compare individual descriptors,

$$C_{\alpha,\beta} \equiv C(\alpha,\beta) = \frac{1}{2}\sum_1^K \frac{[h_i(k)-h_j(k)]^2}{h_i(k)+h_j(k)} \qquad (2)$$

Where $h_i(k)$ and $h_j(k)$ denote the K-bin normalized

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:6, 2011

histograms at point α and β, respectively; $C_{a,b}$ is the distance between the two points.

In the verification stage, given a hand image, we get the key points and their shape context descriptor as described in pre-phrase. The distances between the key point and the k cluster centers are calculated. The final matching score is given as the ratio between the best and the mean similarities, as in (3). If the r is less than threshold, we discriminate the point is true, else discard it.

$$r = d_{min}/d_{mean} \qquad (3)$$

B. ROI acquisition

Use the information of the key points on the hand contour; we can locate a stable ROI area. In [13], Chin-Chuan Han et al. use the key points between the index finger, the little finger, the middle finger and the ring finger to get a ROI region of fixed size. Their method is not applicable under the condition of various palm scale. We improve their approach, use the mean distance between all hand key point pairs to respect the scale of palm. Then we can get a ROI of variable size depend on the palm scale, which is robust to the different poses and scales.

As Fig.5 shows, according to the position of P^1、 P^2、 P^3、 P^4 we can judge whether it is a the right hand or a left hand. Then the key point C^1 between middle finger and index finger, the key point C^2 between middle finger and ring finger, and the key point C^3 between ring finger and pinky finger are located. The steps on obtaining the square area ROI are described as follows:

1) Line up C^1 and C^3 to get the X-axis of image and then make a line through C^2, perpendicular to the Y-axis, and the intersection is d^1.

2)Calculate the mean distance δ between the point C^1, C^2, C^3, and make

$$d^1 d^2 = \frac{\delta}{2}$$

3) Locate the square $a^1 a^2 a^3 a^4$, $a^1 a^2$ and $C^1 C^3$ are parallel, make

$$a^1 a^2 = \delta \times \frac{3}{2}, \qquad \frac{a^1 d^2}{d^2 a^2} = \frac{C^1 C^2}{C^2 C^3}$$

4) Get the ROI image from the original image, and transform it into a gray image of fit size.
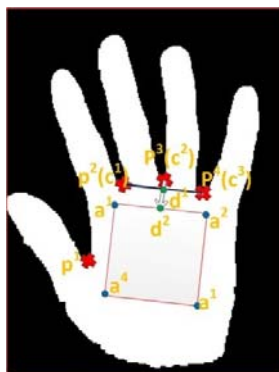

Fig.5 Locate the ROI of palm with key points

VI. EXPERIMENT

For the purpose of training, testing and evaluation of the proposed system, we have captured hand images from 54 individuals with a web camera, whose age ranges from 20 to above 50. The resolution of the image is 640 480 (in pixels) and 24-bit colors. Every people's right hand and left hand were captured in three senses of different light and background. Each hand is collected 20 images in a sense of different pose, such as close-hand, open hand, and half open hand etc. There are totally 6480 hand images in our database. We divide them equally to test set and training set. The experiments are used a common PC with Intel Pentium(R) D 6.0 Ghz and 2GB RAM, implemented with Microsoft Visual C++ 6.0.

A. Detect target area

To train the boosted classifier cascade, we used 1620 right hand images from test set as positive examples, and negative examples were extracted from a set of background images with different scene. Then we used Lienhart and colleagues' approach, implemented in OpenCV. The other 1620 left hands were done in the same way. During detection, rather than scanning in all scales in [10], we just search across some certain scales in which we are interested. Some false positives can be discarded after skin segmentation. We use 2000 hand images from training set, and 500 images without hand from internet as test set. Our final target detector detects 1887 hands in the test set, with 81 false positives. The whole target detection can be finished within 40 ms. Table I compares the performance of the original boosted classifier cascade of Lienhart and Maydt to that of our target detector system which use the scale and color information.

TABLE I.
Performance comparison the result on the ROC curve between a system with only the boosted classifier cascade and our target detector system in terms of the number of false positive detections. Six results from the ROC curve were selected for the comparison.

Exprient Number	Number of true positive from system with only boosted classifier cascade	Number of false positive from system with only boosted classifier cascade	Number of false positive from our system
1	1830	35	20
2	1852	57	31
3	1863	92	49
4	1886	113	61
5	1895	237	105
6	1925	426	187

B. Key point detect and verification

We first test our key point detection algorithm. 2000 binary hand images processed by the target area detection method are used to test. Every hand image should be located by four key points between the five fingers. Our algorithm detects 3.9 true positive points and 1.12 false positive points per image. The average time of processing one image is 25 ms.

In the second experiment, we use another 1000 contour images different from the learning stage which manually

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:6, 2011

marked with true key points. 500 negative points are selected from the original edge map. Let us take a look at the verify results. Fig.6 plots the ROC curves for different values of k with the value of r changed. Table 2 show respectively the average time in verify all key points in an image and the evaluation measures for different values of k. In our system, we chose k = 20, r = 0.35 with true positive rate equaling to 97.3% and false positive rate equaling to 1.7%.
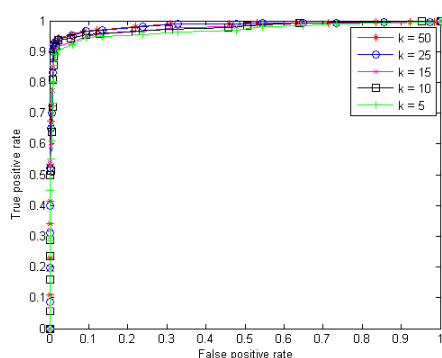
Fig.6 ROC curves for different values of k

C. Evaluation of our system

To measure the accuracy of the system, we test 3240 original images in the test set. If all of the three key points used to locate the ROI are detect accurately, it will be consider as a successful ROI localization. Fig.7 shows some result of our experiment. In the third row of Fig.7 (a), the people wearing a ring or a bracelet also can use our system. The accuracy rate of our system is 93.8% and the average time of processing one image is 178 ms.
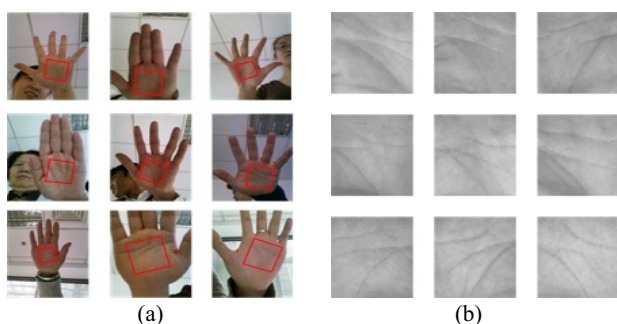

Fig.7 (a) Some results of ROI localization; (b) The ROI of palmprint acquired with our system .

VII. CONCLUSION

In this paper, we developed a novel system to acquire ROI of palmprint in unconstrained scenes. It is robust to the changes of hand pose and work well under cluster background. It process the image captured with a web camera in real-time and achieves promising result.

In the future we will use this preprocess system and develop a palm print recognition system with no restriction on palm,

which will make users more comfortable.

REFERENCES

[1] D. Zhang, W.K. Kong, J. You, and M. Wong, "On-line palmprint identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1041–1050, 2003.
[2] Chin-chuan Han, "A hand-based personal authentication using a coarse-to-fine strategy," *Image and Vision Computing*, vol. 22, no. 11, pp. 909-918, 2004.
[3] Julien Doublet, Olivier Lepetit and Marinette Revenu, "Contact less Hand Recognition Using Shape and Texture Features," in *Proc. 8th International Conference on Signal Processing*, Beijing, 2006, pp. 26-29.
[4] Julien Doublet, Marinette Revenu and Olivier Lepetit, "Robust GrayScale Distribution Estimation for Contactless Palmprint Recognition," in *IEEE Conference on Biometrics: Theory, Applications and Systems*, Washington DC, 2007, pp.69-75.
[5] J. Doublet, Lepetit M. Revenu, "Contactless Hand Recoginition Based On Distribution Estimation," in *Proc. 6th Biometrics Symposium,* Florida, 2008, pp. 69-74
[6] Ivan Fratric, Slobodan Ribaric, "Real-Time Model-Based Hand Localization for Unsupervised Palmar Image Acquisition," in *International Conference on Biometrics*, Sardinia Island, 2009, pp. 1280-1289.
[7] Goh Kah Ong Michael , Tee Connie , and Andrew Beng Jin Teoh, "Touch-less palm print biometrics: Novel design and implementation," *Image and Vision Computing* , vol. 26, no. 12, pp. 1551-1560, 2008.
[8] Michal Choras, and Rafal Kozik, "Feature Extraction Method for Contactless Palmprint Biometrics," in *International Conference on Intelligent Computing*, Changsha, 2010, pp. 435-442.
[9] Yufei Han, Zhenan Sun, Fei Wang, and Tieniu Tan "Palmprint Recognition Under Unconstrained Sce*nes,"* in *Asian Conference on Computer Vision,* Tokyo, 2007, pp. 1-11.
[10] Paul Viola, and Michael Jones, "Robust Real-Time Object Detection," *Journal of Computer Vision,* vol. 57, no. 2, pp.26-33, 2004.
[11] Son Lam Phung, Abdesselam Bouzerdoum, and Douglas Chai, "Skin Segmentation Using Color Pixel Classification: Analysis and Comparison," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 1, pp. 148-154, 2005.
[12] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509﹣522, 2002.