# A Novel Vision based Finger-writing Character Recognition System*

Lianwen Jin[1], Duanduan Yang[1], Li-Xin Zhen[2], Jian-Cheng Huang[2]
[1]School of Electronic and Information Engineering, South China University of Technology, Guangzhou, 510640, China
[2] Motorola China Research Center, Shanghai, 210000, China
E-mail: {eelwjin,ddyang}@scut.edu.cn , {Li-Xin.Zhen, Jian-Cheng.Huang}@motorola.com

## Abstract

*A new vision based finger writing character recognition system (FWCRS) is proposed in this paper. The FWCRS allows people to write characters virtually just using his finger-tip (we call this "finger-writing"). The trajectories of the finger-tip are tracked and reconstructed as a kind of inkless character pattern and finally recognized by a classifier. In this paper, a simple but effective background model is built for the FWCRS to segment human finger from cluttered background. A robust fingertip detection algorithm based on feature matching is presented. The finger-writing character is finally recognized by a DTW classifier. Experiments show that the FWCRS can recognize finger-writing uppercase & lowercase English characters with the accuracy of 95.6%, 98.5% respectively.*

## 1. Introduction

Vision-based HCI (Human Computer Interaction) is an important technology to make machine more intelligent [1,6]. One type of vision-based HCI application is based on tracking and detection of finger or fingertip, such as gesture recognition[6], sign language recognition[5], finger mouse[2][3], augmented desk interface [4][9], finger painting [3] and so on. The fingertip could also be used to write character virtually. Through computer vision and pattern recognition technologies, it is possible for computer to recognize the finger-writing characters.

In this paper, we propose a novel finger-writing character recognition system (FWCRS). The basic idea of FWCRS is that people can write characters virtually with the movement of finger-tip (we call this "finger-writing"). The trajectories of the finger-tip is detected and tracked real timely, and then reconstructed as a kind of inkless character pattern. The reconstructed character is finally be recognized by a classifier to give the output. The diagram of the proposed FWCRS is shown in figure 1.

As the digital camera becomes more and more popular in many portable devices, character inputting based on camera has a significant role to play in the development of novel HCI interfaces, because cameras can be miniaturized thus making the interface much smaller and convenient.

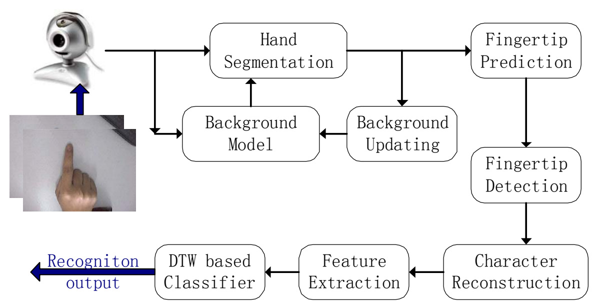



**Fig 1. Diagram of the FWCR system.**

## 2. Fingertip segmentation from cluttered background

Background modeling is one important computer vision problem. Many methods have been proposed to address some of the solutions (eg. [3,7,8]). For our FWCR system, we proposed a simple but effective background model based on the similar idea of W4[8]. Our background model is given by:

$$\mathrm{M}_t(\mathrm{x}) = \begin{bmatrix} \mathrm{m}_t(x) \\ d^t_{median} \\ d^t_{\max} \end{bmatrix} = \begin{bmatrix} \mathrm{m}_t(x) \\ median_z\left\{\left|I^t(\mathrm{x}) - I^{t-1}(\mathrm{x})\right|\right\} \\ \max_z\left\{\left|I^t(\mathrm{x}) - I^{t-1}(\mathrm{x})\right|\right\} \end{bmatrix} \quad (1)$$

where $x$ is pixel index over the current frame. $I_t(x)$ is the intensity of pixel $x$ at frame $t$. $\mathrm{m}_t(\mathrm{x})$ is the mean of background intensity of pixel $x$ over several frames.

After $\mathrm{M}_t(\mathrm{x})$ is determined, the following equation is used to judge if a pixel $x$ belongs to foreground or not:

*This paper is partially sponsored by the following foundations: Motorola Labs Research Project, New Century Excellent Talent Program of MOE (No.NCET-05-0736), NSFGD(No. 04105938)*

0-7695-2521-0/06/$20.00 (c) 2006 IEEE

$$B(x)=\begin{cases}0 & background\cdot if \quad \begin{array}{l} I^t(x)-m(x)<k_{median}\cdot d^t{}_{median} \\ \vee\, I^t(x)-m(x)<T \quad \vee\, I^t(x)-m(x)<k_{\max}\cdot d^t{}_{\max}\end{array} \\ 1 & foreground \quad otherwise\end{cases} \tag{2}$$

In equation (2), $k_{median}$ (typically 1.6), $k_{max}$ (typically 0.25) and T(typically 25) are three constants.

Experiments showed the background model based on eq.(1)~eq.(2) worked well in indoor situation. Figure 2 shows some experimental results of our model and W4 model [8]. In figure 2, column 1 from left shows a normal situation. Column 2 & 3 are the situations when a flash light was shining by. Column 4 and 5 gives some situations where intensity of some background objects becomes similar with foreground. It can be seen that the proposed background model can adapt to light intensity changes more effectively.

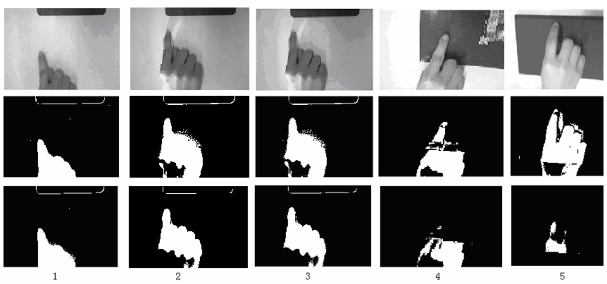



**Fig.2 Experimental results of our model and W4 model. The 1st row shows origin pictures, the 2nd row gives foregrounds extracted by our background model, and the 3rd row gives results of W4.**

## 3. Fingertip Detection

Previous fingertip localization methods mainly include contour analysis[2][3], template matching[3,4], heuristics method[9], and so on. However, experiments shows most of the methods tend to fail under the condition that the foreground is not segmented clearly. Some of these approach can only works well under some special constraints, such as white background, clear contrast between hand and background, hand motion cannot be very fast, the direction of fingertip must be up, etc. To overcome these constraints, we propose a new fingertip detection approach. The diagram of our fingertip detection approach is shown in figure 3.

### 3.1 Rough localization of fingertip

Hand image is usually used to analyze the fingertip position. However, if we use of the whole hand image to analyze the position of fingertip directly, too much computation will be involved. So we propose a rough fingertip localization method based on down-samples of the contour of the hand. As shown in figure 4, firstly we got the hand contour from segmented foreground (with size of 320×240) using edge operator (see figure 4(c) ). Then a grids of 10×10 is marked for the contour image (figure 4(d)). Each grid of the contour image will be finally mapped to a while or black pixel according to whether there are edge pixels in each grid or not. By this way, the original 320×240 hand contour is been mapped to a down-sampled 32×24 contour, so that much less computation will be involved in the following fingertip detecting stage. The down-sampling can be also regarded as a kind of low-pass filter, therefore high frequency blur of hand contour caused by fast motion can be removed.

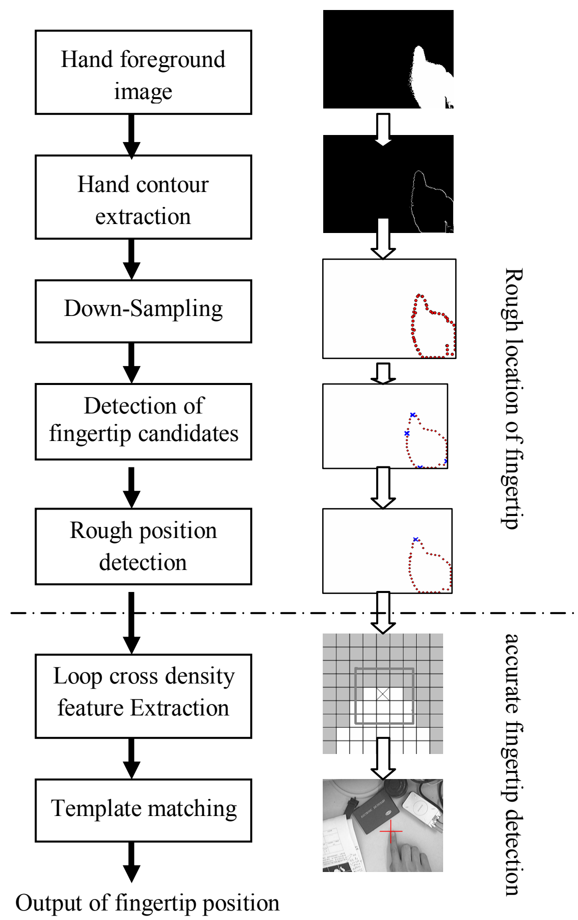



**Fig3. Diagram of fingertip detection.**

In our FWCRS, it is found the rough position of the fingertip must be one of four directional peaks of the close curve of the finger contour. These four peaks are treated as the candidates for the rough position of the fingertip (see figure 4(e)). The overall shape of a human finger can be approximated by a cylinder with a hemispherical cap. We select several sample point-pairs starting from each candidate peak anticlockwise and clockwise. The variance of the distances of these point-pairs is calculated. The peak corresponding to the minimal variance is regarded rough fingertip position. (see figure 4(f)).

### 3.2 Accurate detection of fingertip

It should be noticed that the fingertip position detected by the method mentioned in section 3.1 is not accurate enough, since the hand contour may not be accurate due to disturb of shadow, or incorrect segmentation of correct hand foreground. So, we need to detection the precise fingertip position from the original image. In real application condition, the moving fingertip may points to different directions in different frames. To solve the rotation-invariant problem, we propose a feature matching method. As shown in figure 5, a set of rectangle (typically 12~18) loops are drawn around the fingertip, the number of pixel density which each rectangle across by is accumulated as a feature. All features will finally form a feature vector. We call this as loop cross density feature **(LCDF)** vector. For example, the LCDF along the three rectangles shown in figure 5 is 5,7 and 7 respectively. To find the precise location of the fingertip, the LCDF vector of fingertip template is matched with an unknown image around the region of rough fingertip position. The position that produces the highest matching score will be regarded as the fingertip position.

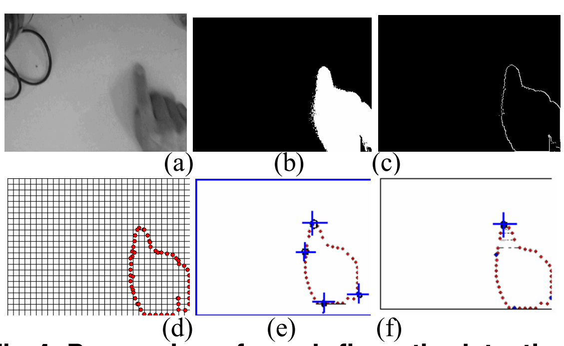



**Fig 4. Processing of rough fingertip detection. (a).original image (b) segmented hand foreground. (c). Contour of hand. (d). Down-samping of hand contour. (e) Detect four peaks form the down-sampled contour (f) Rough fingertip detection**

Figure 6 gives a comparison of the fingertip detection results using traditional template matching method and our LCDF matching method. It can be seen that when the orientations of fingertip is not pointed up (as that of the template), traditional template method is not good enough. The proposed feature matching method is robust when the fingertip points to different orientations. Figure 7 gives more experimental results.

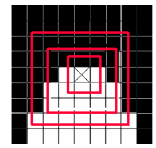


**Fig 5. Extraction of loop cross density feature**

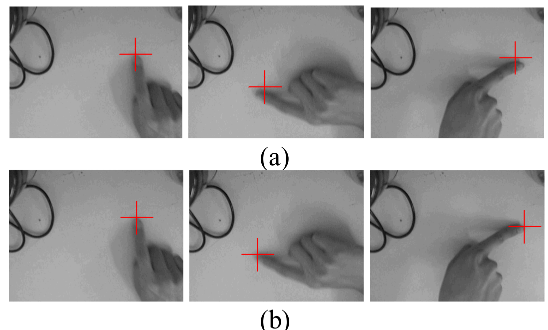



**Fig 6. (a) Fingertip detection results using traditional template matching method. (b) Results using the proposed loop cross density -feature matching method.**

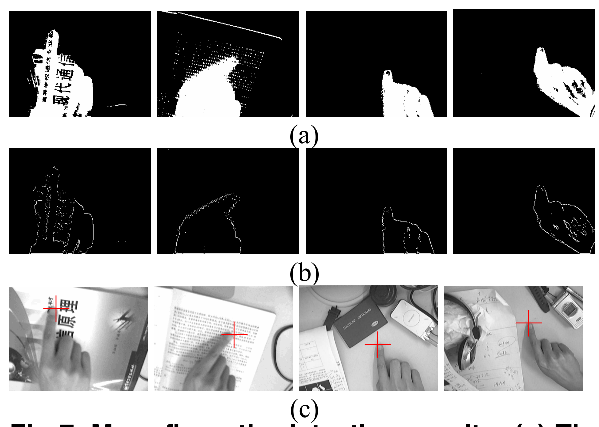



**Fig 7. More fingertip detection results: (a) The segmented foreground. (b) The contour of the hand (c) The localization of fingertip in corresponding original images.**

## 4. Finger-writing Character Recognition

### 4.1 Reconstruction of finger-writing character

The reconstruction of finger writing characters is implemented by linking all detected consecutive fingertip positions together. However, since the trajectories of the fingertip are inkless, one problem is how to tell which trajectory is the beginning or ending writing point. Based on experimental observation, we using the following three conditions to judge the start writing point: (1). The fingertip position is in the region from the top to the middle line; and (2). the vertical or horizon displacement is more than 15 pixels in 3 consecutive frames; and (3). the angle between the two vectors composed by the fingertip position in 3 consecutive frames is less than a given threshold. Similarly, if either of the following two conditions is satisfied, the detected fingertip trajectory is regarded as the end writing point: (a). Pixels belonging to the foreground less than threshold; or (b). the position of fingertip keeps in the same place for 15 frames. Some examples of the reconstructed characters are given in figure 8.

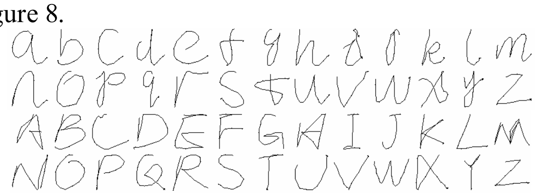


**Figure 8. Reconstructed finger-writing character samples**

### 4.2 Recognition of Finger-writing Character

On special characteristic of the reconstructed character is that all stroke of it are connected (since it is difficulty to distinguish the state of up and down of fingertip when writing with finger). So it is a kind of one-stroke style character. We use a DTW (dynamic time warping) classifier with combination of directional and positional features for the recognition of the characters [10]. The classifier is based on prototype matching by DTW with stroke directional and positional features. More detail could be found in [10].

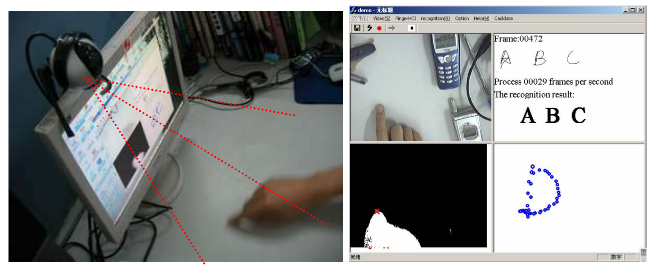
### 4.3. Experimental results

We collected 69 sets of uppercase and lowercase English letters written by different persons using our FWCRS. We use the templates constructed by hand to test the efficiency of finger-writing English character recognition. The uppercase letter templates consist of 151 prototypes, lowercase letter templates consist of 258 prototypes. The size of the template file is 3.86K and 4.88K respectively. Experiment on the 69 sets of uppercase and lowercase English produces the recognition accuracy of 95.6%, 98.5% respectively.

The proposed FWCRS system has been implemented on a Pentium PC of 3.0G CPU with 512M memory. We use an ordinary low resolution USB digital camera to capture the finger movement. The resolution of image is $320 \times 240$. The speed is about 28~30fps (depending on the complexity of the background), showing it is suitable for real time application. Figure 9 illustrates our FWCRS.

## 5. Summary

A novel vision-based finger-writing character recognition system (FWCRS) is proposed in this paper. With the FWCRS, human can input character to computer just using the movement of his fingertip, without any additional device such as keyboard, touch screen or digital pen. Compare with conventional character inputting modalities, the proposed FWCRS can make the device much smaller without loss of efficiency. It provides a novel interesting character inputting modality.




(a) (b)

**Fig 9. (a). System setup. (b). User interface of the FWCRS.**

## Reference

[1] V. Pavlovic, R. Sharma etal., Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review, *IEEE* Transactions on PAMI, *Vol. 19, No. 7, 677-695，1997.*

[2] F. Quek, et.al. *Finger mouse: A freehand pointing interface*. Proc. of International Workshop on Automatic Face and Gesture Recognition. 1995. Zurich, Switzerland.

[3] C. Hardenberg, François Bérard, Bare-Hand Human Computer Interaction, Workshop on Perceptive User Interfaces, Florida, USA, Nov,15-16, 2001

[4] J. Crowley, F. Bérard,. Finger tacking as an input device for augmented reality, International Workshop on Automatic Face and Gesture Recognition", Zürich, 195-200, 1995.

[5] K. Imagawa, et.al, Color-Based Hands Tracking System for Sign Language Recognition, International Conf. on AFGR98, pp. 462--467, 1998.

[6] Kenji Oka, Yoichi Sato, Hideki kOike, Real-time Tracking and Gesture Recognition,IEEE Trans. Computer Graphics and Application, pp. 64 -71December 2002.

[7] A. Elgammal, R.Duraiswami, D.Harwood and L.S.Davis, *Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance*, Procceedings of the IEEE, Vol.90, No.7, 2002

[8] I. Haritaoglu, et.al., *W4: Real-Time Surveillance of People and their Activities*,IEEE Trans. On Pattern analysis and machine intelligence, Vol.22, no.8, pp809-830, 2000.

[9] Brown.T, et.al, Finger tracking for the Digital Desk**,** First Australasian User Interface Conference, pp11-16, 2000.

[10]Teng Long, Lianwen Jin, et al., One stroke cursive character recognition using combination of directional and positional features, Proceedings of IEEE ICASSP, pp449-452, 2005.