AN EFFECTIVE ROBUST FINGERTIP DETECTION METHOD FOR FINGER WRITING CHARACTER RECOGNITION SYSTEM

DUAN-DUAN YANG[1], LIAN-WEN JIN[1], JUN-XUN YIN[1], LI-XIN ZHEN[2], JIAN-CHENG HUANG[2]

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

Abstract:

This paper proposes an effective and robust fingertip detection method in 2-D plane and applies it to a novel vision based human computer interaction system: Finger Writing Character Recognition System (FWCRS). The fingertip detection approach consists of two stages. First, based on the grid sampling and the analysis of sampled hand contour, the fingertip was detected roughly. Then, the location of fingertip was localized precisely based on circle feature matching. Experiments suggest that the proposed fingertip detection method is capable of detecting fingertip in a reliable manner even in a complex background under different light conditions without any markers. To demonstrate the strength of the method, the method was run on 5 sequences with varying light condition, different degrees of clutter background and different speeds of finger movement, experiment shows that the correct rate can reach 98.5%.

The finger writing character recognition system in this paper is particularly advantageous for Human-computer Interaction (HCI) in that users can communicate with computers by their favorite mean: handwriting. At the same time, they can perform handwriting with only their finger directly.

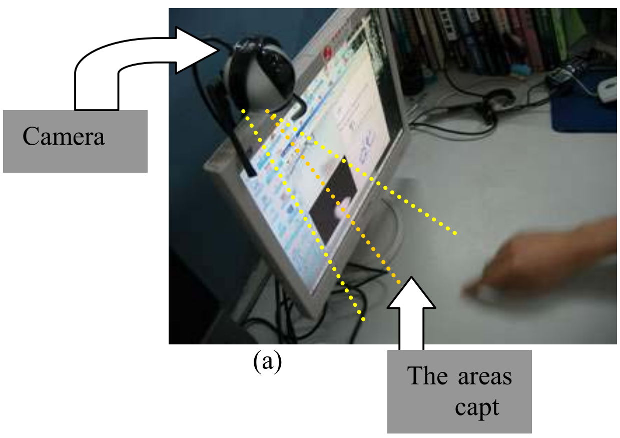
Keywords:

Fingertip detection; background; template matching; circle features

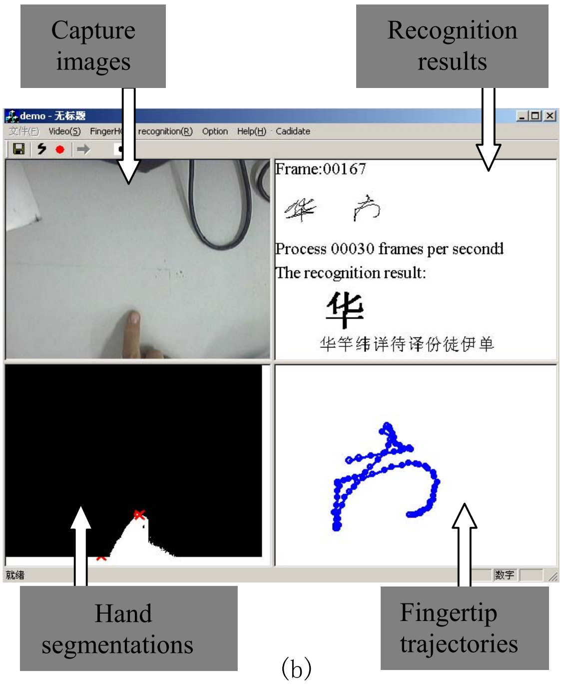
1. Introduction

In the artificial intelligence field, people are trying to develop some intelligent vision-based interaction systems those are more intuitive than traditional interaction devices such as mice, keyboard. Tracking fingertip in 2-D plane as the input for the computer [1-5] is one desirable mode in human-computer interaction. Inspired by those system, we developed a finger writing character recognition system (FWCRS), shown in the Figure 1(a), whose interface was shown in Figure 1 (b). There are four sub-windows, they show captured images, hand segmentation results, fingertip trajectories and the recognition results. The advantage of our character recognition system is that it enables users can input characters to the computer by handwriting with their own fingers.

The key component of FWCRS was a vision-based method for tracking user' fingertip in real time. Towards this goal, fingertip detection method should be fast and robust. Almost all of fingertip detection methods are based on hand segmentation technique because it can decrease the amount of image information by selecting areas of interests for fingertip detection. However most hand segmentation methods could not provide a clearly hand segmentation under some conditions (fast hand motion, cluttered background, poor light condition)[4]. Poor hand segmentation performance usually invalidates fingertip detection methods. The methods [5,9,10] made uses of infrared camera to get a reliable segmentation, the methods in [1,2,3,4,11,12] limit the degree of the background clutter, finger motion speed or light conditions to get a reliable segmentation. On the other hand, some of previous fingertip detection methods cannot work accurately when they come to localizing multi-direction fingertips. To get a precise localization, they usually [1,2,8,11] suppose the finger always point upward.



(a)



(b)

Figure 1. Overview of FWCRS. (a). System setup. (b). The interface of our FWCRS.

In this paper, a fast and robust fingertip detection method is proposed that can localize multi-direction fingertips precisely even without a clear hand segmentation results. First, based on the analysis of the samples of hand contour, the candidate region for fingertip localization was selected. Then, the location of the fingertip was located based on circle feature matching and the region.

To demonstrate the strength of the method, we ran the method on 5 sequences with various light condition, different degrees of background clutter and different speeds of movement, experiment shows that the correct rate of this fingertip detection method can reach 98.5%. In addiction, the effectiveness has been successfully demonstrated by the proposed finger writing character recognition system.

2. Brief Summary of Previous work

Previous fingertip localization methods in 2-D plane mainly include the three categories: contour analysis [2,3,4,8,12], template match [3,4,5] and heuristics methods [11]. The effectiveness of all of the previous methods depends on the performance of hand segmentation. To get reliable segmentation, they tend to limit the degree of the background clutter, finger motion speed or light conditions, or using infrared camera to capture images.

Table 1 summaries some typical fingertip detection approaches and their limitation.

Table1. Some typical previous fingertip detection approaches and the conditions assure the methods success

Apply	Method	Condition
[1]	Template matching with a fingertip template	The illumination of the workspace is controlled and generally uniform.
[2]	Axis-boundary intersection	Single background: keyboard
[3]	Conic fitting	Hand color is distinguishable from the panel's.
[4]	Based on finger shape	A fairly clean segmented region of interest
[5][9][10]	Template matching	Using the infrared camera to capture images
[12]	Contour tracking	A clear contrast between the hand and background
[11]	The top of foreground pixels was assumed to be the fingertip	White background

3. Fingertip detection method

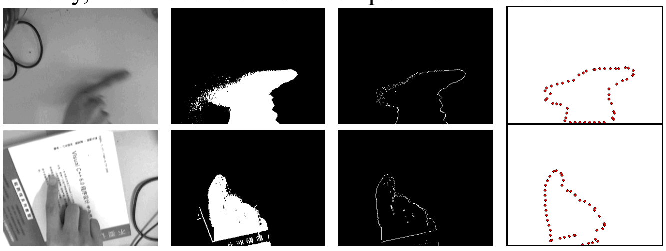
Like previous fingertip detection methods, we perform hand segmentation before localizing the fingertip. In this paper, we use the image differencing technique [4] to segment the 'hand'.

The fingertip detection is divided to two stages to accomplish. First, the rough position of the fingertip is found out by analyzing the samples of hand contour. Based on the rough position, the candidate region was selected for precise fingertip localization. Second, *circle feature matching* was used to localize fingertip precisely. As finger writing is complex course, prediction technique (e.g. Kalman Filter) could not get a satisfied result, we don't plan to involve the prediction in this paper.

3.1. Rough localization of fingertip

To get the hand contour, we use the grads operator to detect the edge in the hand segmentation images (e.g. Figure. 2 (b).

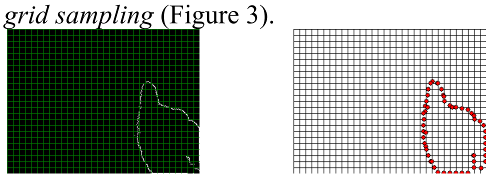
As it can be seen, under some condition (fast hand motion, cluttered background, poor light condition), the segmentation method cannot provide a clearly hand segmentation (e.g. Figure. 2 (b)). As a result, hand contours maybe are blurred and distorted (e.g. Figure. 2(c)). If those hand contours are used to analyze the position of fingertip directly, it will cost too much computation and tend to fail.



(a) Captured images (b) Hand segmentation (c) Hand contours (d) Sampled hand contours

Figure 2. Some examples of grid sampling

In this paper, we search a rough fingertip position based on the analysis of the samples of hand contours. Supposing the size of video images is captured as 320*240 pixels. Fist, making the grids (10*10 pixels) for hand contour images (edge images, e.g. Figure 3(a)). Second, sampling hand contours by the girds according the following mapping rules: Each gird in hand contour images (e.g. Figure 3(a)) is mapped a pixel in sampled contour images (e.g. Figure 3(b)) which is set to black if there are edge pixels in the grid, other is set to white. We call the course as <u>*grid sampling*</u> (Figure 3).



(a) Hand contour images (b) sampled hand contours

Figure 3. Grid sampling

By *gird sampling*, some blurred or distored contour (e.g. figure2 (c)) become clear and smoothly without lost useful fingertip information(e.g. figure2 (d)). Experiments show the analysis of the samples of hand contours is more reliable to fingertip localization than the analysis of hand contours themselves. By *gird sampling*, we get a sampled hand contours images with low-resolution images (32*24). Obviously, analyzing it involve less computation than analyzing the contours images.

In sampled hand contours images, the longest close curve defined as *sampled hand contours*. Based on many experiments results, we found the rough position of the pointing fingertip must be one of peaks of the *sampled hand contours*. So the four peaks of a *sampled hand contour* are chosen as the candidates of the rough position of the fingertip. Based on the facts that the overall shape of a human finger can be approximated by a cylinder with a hemispherical cap and the width of the cylinder is almost same to different people, around each candidate, we respectively select 4 points on the *sampled hand contours* anticlockwise and clockwise to make up 4 points-pairs. The variance of the distances of those points-pairs around each candidate is calculated. The candidate corresponding to the minimal variance is regarded as rough fingertip position.

In summary of above description, we use figure 4 to show our course of rough localization algorithm.

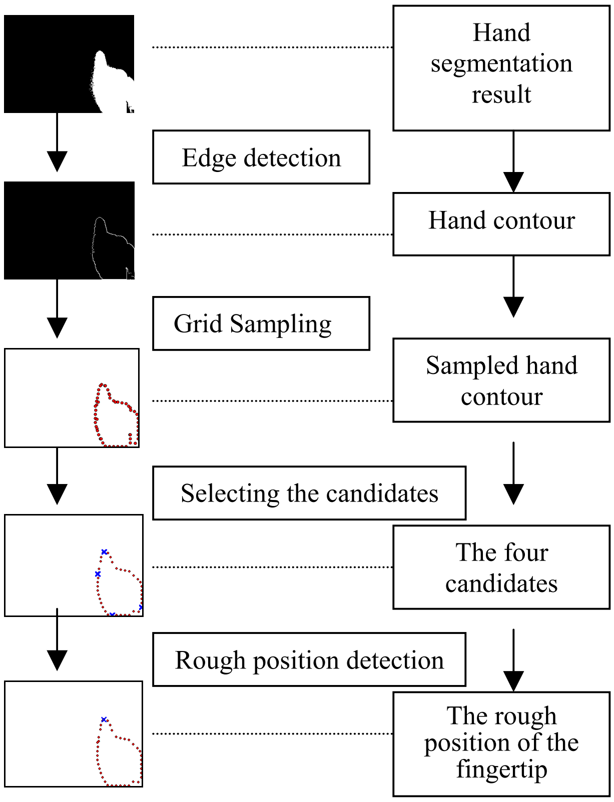


Figure 4. Rough localization of fingertip

3.2. Precise localization of fingertip

In many finger interface systems, the finger points in different direction in different frames. To multi-direction fingertips localization, experiments show traditional template matching could not provide enough precise localization. Multi-template matching can solve the problem but it will bring some other problems such as template choosing and computation complexity at the same time. A circular template are used in [5] can localize the fingertip that points in different direction, but too much candidates are detected. So two means are used to remove the wrong detection in [5].

Inspired the circular template, we propose the conception of *circular features* that is steady when the finger pointing in different direction. Supposing there is a binary image, where white pixels denote foreground and black pixels denote background, we want to calculate the *circle features* of the pixels: (i_0, j_0) of the image. First, Setting the dimension of circle features: k . Second, drawing k squares with different sizes around ((i_0, j_0) and make sure all of those pixels passed by the sides of the kth squares have same chessboard distance to (i_0, j_0): k pixels. At last, counting the number of white pixels those was passed by the kth square as the kth dimension of *circle feature*. For example, we calculate the circular feature of the '×' pixel in figure 5, the 2nd dimension of the circular feature of the '×' pixel F_2 is 7.

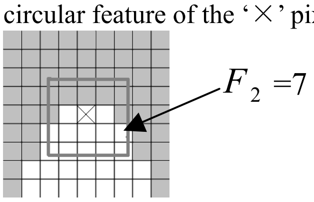


Figure 5. An example of circle feature

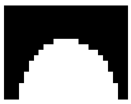
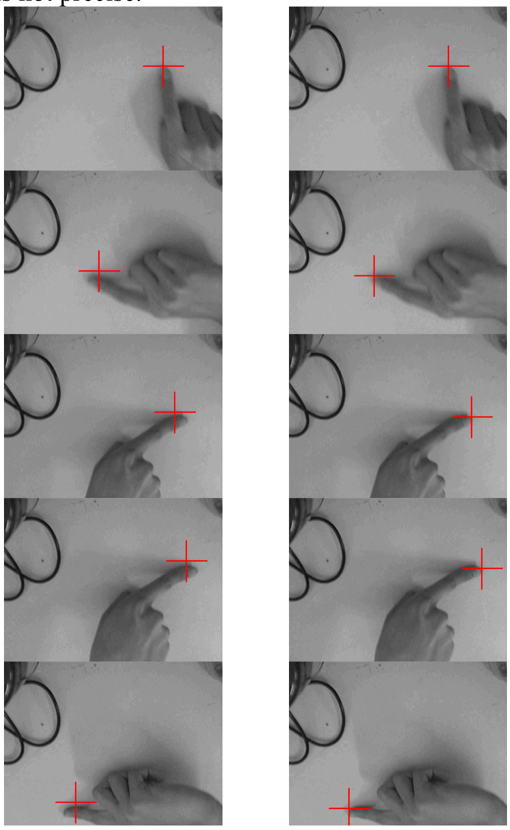


Figure 6. The template

In practice, the circular feature of (i_0, j_0) can be calculated as $F(k) = \sum_{Dc((i,j),(i_0,j_0))=k} S(i,j)$, where F denote the feature and $S(i,j) = \begin{cases} 0 \ldots\ldots (i,j) \notin foreground \\ 1 \ldots\ldots (i,j) \in foreground \end{cases}$, $Dc((i,j),(i_0,j_0))$ denotes the chessboard distance between (i,j) and (i_0, j_0), k is the dimension of circle features, (i,j) denote the pixels around (i_0, j_0).

In this paper, k is set to 12. Then, we can use the method of circle feature matching to detect fingertip. First, We can center the rough localization of the fingertip to define the candidate region whose size is about 10 pixels *10 pixels for *circle feature* matching. Second, we calculate the *circle features* of every edge pixels in the region and the template (Figure 6). Third, by the feature matching, the matching score is highest is considered the fingertip.

The Figure 7 gives the comparison of the traditional template matching and the circle feature template matching, both of them are using the template to match the foreground binary image. As we can see, when the direction of fingertip is not upward, the traditional template matching method is not precise.



(a) (b)

Figure 7. A comparison of the traditional template matching and our method, (a) the results from the traditional template match, (b) the results from the circle-feature template match.

4. Performance comparison with previous methods

As the description above, searching region in rough localization is the samples of hand contours, which usually less than 100 pixels. In precise localization, searching region is the defined feature matching area: 10 *10 pixels. In another word, the searching region of our method about is 100 +10*10=200 pixels. The total latency of the method was between 20 and 26ms depending the number of the samples of the hand on a personal computer with PIII 866 MHz

Table 2 shows the time-consuming comparison between previous methods and our method. From it we can see our method outperform the previous methods in speed aspect.

Table 3 shows the comparison between previous methods and our method in the aspects of limiting the degree of the background clutter, finger motion speed, light conditions and the finger pointing direction. From it we can see our method is more robust than previous methods.

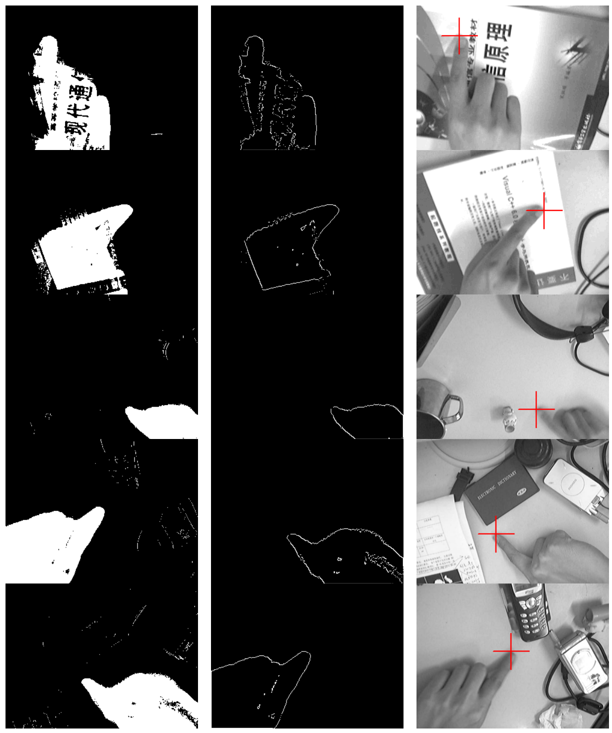
Figure 8 shows some fingertip results of our algorithm.

Table 2. Time-consuming comparison with previous methods

Apply	Image sizes	Searching region sizes	Speed/ time consumption
Visual panel [3]	320*240	About 30*30	Real time
HCI [4]	384*288	Foreground pixels	26-34ms
Gesture recognition [9]	256*220	80*80	25-30fps
Input device [1]	192*144	About 26*26	24fps
Finger mouse [2]	-----	Foreground pixels	7fps
Gesture interface [11]	-----	Foreground pixels	Not real time
Computer interface [12]	-----	The pixels of the contour	Real-time
Digital desk [8]	512*512	About 40*40	Real time
Our method	320*240	Most less than 20*20	20-26ms

Table 3. The Comparison with previous methods in some restrictive conditions

Apply	Background	Pointing direction	Motion speed	Remark
Visual panel[3]	Clutter	Random	------	Distinguishable color
HCI [4]	Clutter	Random	<4.5m/s	A fairly clean segmentation
Gesture recognition [9]	Clutter	Random	Random	Infrared camera
Input device [1]	Simple	Nearly up	<1.39m/s	Uniform illumination
Finger mouse [2]	Keyboard	Up	------	Single background: keyboard
Gesture interface [11]	White	Up	-----	White background
Computer interface [12]	Clutter	Random	-----	Analyzing hand contour s
Digital desk [8]	A clear contract	Up	< 1.39m/s	A clear contrast
Our method	Clutter	Random	Random	-----------



(a) (b) (c)

Figure 8. some results of our method: (a) hand segmentation results (b) hand contours (c) fingertip localization results

5. Conclusions

This paper proposed a robust and effective fingertip detection method base on the analysis of the samples of hand contours and the circle feature matching. The fingertip detection method is capable of detecting fingertip in a reliable manner even in a complex background under different light conditions, different scenes without any markers. It can also solve the motion blur problem very well by proposed grid sampling technique. To test accuracy and robustness, we ran the finger localization algorithm on 5 images sequences with various light conditions and background without. Experiments shows using our method, localization accuracy limiting the speed of hand 'motion' and pointing orientation can reach 98.5%. In addiction, an application system, finger writing character recognition system, can work very well based on the proposed fingertip detection method.

Acknowledgements

This paper is sponsored by: Motorola Labs Research Foundation (No.303D804372), NSFC (No.60275005), GDNSF (No.2003C50101, 04105938).

References

[1] J. L. Crowley, F. Berard and J. Coutaz, "Finger Tacking As an Input Device for Augmented Reality," Proceedings of. Int'l Workshop on Automatic Face and Gesture Recognition, Zurich, Switzerland, pp:195-200, June 1995.

[2] F. K. H. Quek, T. Mysliwiec, and M. Zhao, "Finger Mouse: A Freehand Pointing Computer Interface," Proceedings of Int'l Workshop on Automatic Face and Gesture Recognition, Zurich,Switzerland, pp.:372-377, June 1995.

[3] Ying Wu, Ying Shan, Zhengyou Zhangy, Steven Shafer, "VISUAL PANEL: From an ordinary paper to a wireless and mobile input device," Technical Report, MSR-TR-2000 Microsoft Research Corporation, http://www.research.microsoft.com, October 2000.

[4] Christian. V. H, François. Bm, "Bare-hand human computer interaction," Proceedings of the 2001 workshop on Percetive user interfaces, Orlando, Florida, USA, pp 1-8, Nov 2001.

[5] Oka. K, Sato. Y, Koike. H, "Real-time tracking of multiple fingertips and gesture recognition for augmented desk interface system", Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, pp 411 – 416,May 2002.

[6] J.M. Rehg and T. Kanade, "DigitEyes: Vision-Based Human HandTracking," Technical Report CMU-CS-93-220, School of ComputerScience, Carnegie Mellon Univ, 1993.Proc. InterCHI'93, Apr. 1993.

[7] Andrew Wu, Mubarak Shah, N. da Vitoria Lobo, "A Virtual 3D Blackboard: 3D Finger Tracking using a Single Camera", Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000, Grenoble, France , pp:526, March, 2000.

[8] Brown. T, Thomas. R.C, "Finger tracking for the Digital Desk", Proceedings of First Australasian User Interface Conference, pp11 – 16, 31 Jan.-3 Feb. 2000.

[9] Oka.K, Sato.Y, Koike,.H, "Real-time fingertip tracking and gesture recognition", Computer Graphics and Applications, IEEE , Volume: 22 , Issue: 6 pp:64 – 71, Nov.-Dec. 2002.

[10] Sato.Y, Kobayashi.Y,Koike.H, "Fast tracking of hands and fingertips in infrared images for augmented desk interface", Proceedings. Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp: 462 – 467,28-30 March 2000.

[11] Tomita. A, Jr. Ishii. R, "Hand shape extraction from a sequence of digitized gray-scale images", 20th International Conference on Industrial Electronics, Control and Instrumentation, 1994, IECON '94, vol.3, pp:1925 – 1930, 5-9 Sept.1994.

[12] Keaton.T, Dominguez.S.M, Sayed, A.H,"SNAP&TELL: a multi-modal wearable computer interface for browsing the environment," Proceedings of Sixth International Symposium on Wearable Computers, 2002. (ISWC 2002), pp:75 – 82, 7-10 Oct. 2002.