Recognition of In-air Handwritten Chinese Character Based on Leap Motion Controller

Ning Xu, Weiqiang Wang$^{(\boxtimes)}$, and Xiwen Qu

School of Computer and Control Engineering,
University of Chinese Academy of Sciences, Beijing, China
`wqwang@ict.ac.cn`

Abstract. The three-dimensional interaction has been widely used as a natural and direct way in Human-Computer Interaction (HCI). In this paper, we propose a novel 3D interaction method by recognizing Chinese character written in the air. Firstly, the moving trajectory of fingertip is precisely captured using the Leap Motion Controller. Then, we describe the trajectory by combining the directional feature and direction-change feature. We construct a dataset called IAHCC-UCAS2014, which contains 3755 classes of Chinese characters and each character class has 65 samples. In the evaluation experiments, the proposed method shows promising recognition performance with little increase in computational cost.

Keywords: In-air handwritten Chinese character recognition · Directional feature · Leap Motion Controller

1 Introduction

To achieve high recognition accuracy and system efficiency in online handwritten character recognition (OHCCR) systems, many features and classifiers have been proposed [1], and some satisfying experimental results have been obtained on the existing datasets [2,3]. Recent technologies in vision sensors are capable of capturing 3D finger positions and movements. To chase a more friendly experience of writing, the conception of in-air writing has been proposed and several writing-in-the-air systems [4–6] have been developed, which brings people's writing behavior to the 3D space now.

Feng et al. [4] proposed a finger-writing character recognition system based on the Kinect sensor. By using the depth information and clustering algorithms, the fingertip is located and then the fingertip's trajectory is captured. This approach has gained high tracking accuracy in the dataset including digits and some Chinese characters. In [5], the algorithms for fingertips' detection and tracking are further improved. Jin et al. [6] proposed a digit string recognition method, where the trajectory captured by Kinect is first over-segmented and then recognized by a path-searching algorithm.

The Leap Motion controller is a new generation of 3D interaction sensor which focuses on the interaction by human hands. It can accurately track the

movement of hands and fingertips in a three-dimensional space [7] and it provides application programming interfaces (API) for related interaction. The Leap Motion Controller has been applied into many fields [8–10] and these applications demonstrate the high performance and practical value of it. In our work, we apply the Leap Motion Controller to provide precise and real-time fingertip positions in its 3D workspace. Actually, writing with the Leap Motion Controller is very user-friendly owing to its excellent performance for fingertip detection. In the proposed system to recognize handwritten characters in the air, users can write a Chinese character in the air by moving their fingers relatively fast and fluently.

Compared with traditional OHCCR, in-air handwritten character recognition (IAHCCR) is technically more difficult due to two reasons: First, in-air writing behavior is more likely to be casual which can result in great variation and distortion of the character's structure. Second, there is no pen-up or pen-down information when writing in the air, because the whole character is written by one single stroke. Examples of the handwritten SCUT-COUCH2009 dataset [2] and some in-air written samples from our IAHCC-UCAS2014 dataset are shown in Fig. 1.

To overcome the challenges in in-air handwritten Chinese character recognition, a more robust feature is needed. We exploit the 8-directional feature [11] widely used in online handwritten Chinese character recognition, since the two problems have a lot in common. The 8-directional feature can reflect writing direction for the input Chinese character, and is relatively robust. In [12], a similar directional feature is introduced and in [13] the 8-directional feature is improved. In [14,15] the direction-change feature is proposed and combined with the 4-directional feature. The direction-change feature reflects the direction variation during the writing process. This paper combines the 8-directional feature with the direction-change feature into our recognition system for IAHCCR.
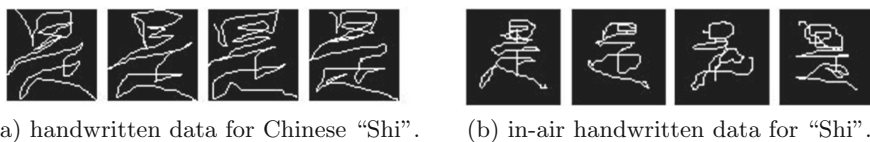

(a) handwritten data for Chinese "Shi".   (b) in-air handwritten data for "Shi".

Fig. 1. Examples of handwritten and in-air handwritten Chinese characters

The rest of this paper is organized as follows: First, we introduce how the trajectory was captured using the Leap Motion Controller. Second, we describe our combined feature using 8-directional feature and direction-change feature. Third, the framework of our recognition system is introduced. Finally, we test the proposed feature on our IAHCC-UCAS2014 dataset and compare the performance with 8-directional feature and the origin direction-change feature.

2   Writing Trajectory Capturing

In our system, by using the Leap Motion Controller, users can move their fingers casually in a customized 3D space. Compared with the Kinect sensor, the Leap Motion Controller has a higher tracking accuracy for fingertips. So the proposed method makes it possible for users to just write by their fingertips with little body movement. Also, owing to the real-time performance of the Leap Motion Controller, users can write relatively fast and naturally, which is hard to achieve in systems based on Kinect sensors. The writing progress by our system can be seen from Fig. 2, where the user is writing the Chinese character of "Shi".
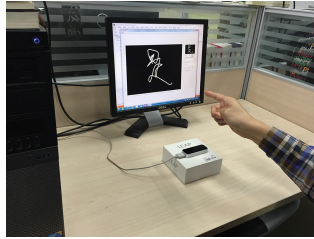

Fig. 2. A user writing a Chinese character "Shi" using our writing-in-the-air system

The 3D writing trajectory can be captured by tracing the movement of the writing fingertip based on the APIs of the Leap Motion Controller. Further, the 2D writing trajectory is obtained by projecting the 3D trajectory onto a screen plane. By adjusting the parameters of Leap Motion Controller, the stability and accuracy of writing trajectory can be guaranteed.

In practice, we find the detection accuracy of the Leap Motion is so high that a slight shake of fingertip can cause apparent structure jitter in the 2D trajectory. So we apply the classic Kalman Filter to smooth the obtained 2D tracking trajectory to reduce the distortion caused by slight shake.

In OHCCR, the imaginary stroke refers to manual straight lines between the end point of one stroke and the start point of its next stroke. In our writing-in-the-air system, the character is always written by one single stroke, so the imaginary stroke is already there for recognition. It should be also noted that the sampling points are usually dense in OHCCR problem, but in our system they can be sparse, since some users can write really fast. So we join the sampling points using Bresenham's line algorithm to construct the final 2D trajectory.

3   Combined Directional Feature

3.1   8-Directional Feature

After several pre-processing steps, each testing sample is normalized into a fixed size of 64×64, and then the 8-directional feature is extracted. Concretely, for a

given point $P_j = (x_j, y_j)$ in the the sequence of sampling points $P_j, j = 1, 2, \cdots$, let $\boldsymbol{V}_j$ denote its direction vector, $\boldsymbol{V}_j$ is defined as follows:

$$\boldsymbol{V}_j = \begin{cases} \overrightarrow{P_j P_{j+1}} & \text{If } P_j \text{ is a start point} \\ \overrightarrow{P_{j-1} P_{j+1}} & \text{If } P_j \text{ is a non-end point} \\ \overrightarrow{P_{j-1} P_j} & \text{If } P_j \text{ is an end point} \end{cases} \qquad (1)$$

(a) Eight direction axes for 8-directional feature.

(b) An mapping example of the direction vector.

Fig. 3. Axes and mapping example of the direction vector

Then its normalized $\boldsymbol{V}_j / \|\boldsymbol{V}_j\|$ is projected to two directions from eight directions as shown in Fig. 3(a). One is from the direction set $\{D1, D3, D5, D7\}$ and denoted by d_j^1, and the other is from set $\{D2, D4, D6, D8\}$ and denoted by d_j^2. Figure 3(b) shows an example, where $d_j^1 = D1$ and $d_j^2 = D8$ for the highlighted sampling point. The corresponding mapping values a_j^1 and a_j^2 for directions d_j^1 and d_j^2 is computed by

$$a_j^1 = \frac{|d_x - d_y|}{s},$$
$$a_j^2 = \frac{\sqrt{2} \cdot \min(d_x, d_y)}{s}, \qquad (2)$$

where $d_x = |x_{j+1} - x_{j-1}|$, $d_y = |y_{j+1} - y_{j-1}|$, and $s = \sqrt{d_x^2 + d_y^2}$ for a non-end point. Further, eight directional pattern images $\{B_d = [f_d(x, y)], x, y = 1, \cdots, 64, d = D1, \cdots, D8\}$ are generated by setting $f_{d_j^1}(x_j, y_j) = a_j^1$ and $f_{d_j^2}(x_j, y_j) = a_j^2$. All the remaining values for $f_d(x, y)$ are set as 0s. The eight directional pattern images are thickened by a maximum filter and then smoothed by a Gaussian filter $G(x, y) = \frac{4}{\lambda^2} \exp[-\frac{2(x^2 + y^2)}{\lambda^2}]$, where λ is the wavelength of the plane wave of the original Gabor filter.

Finally, each directional pattern image is divided uniformly into 8×8 grids. In each grid, the values are summed up to get a feature value. Since we have 8 images and each image has 64 grids, we obtain $8 \times 64 = 512$ dimensional feature vector. A nonlinear transformation (the square root function) is applied to form the final 8-directional feature vector.

3.2   Direction-Change Feature

The direction-change degree and the directions after direction change are obtained from the normalized on-line data by using the direction-change feature. For each sampling point P_j, the direction-change degree is measured by the absolute value of the difference in direction from direction vector $\overrightarrow{P_{j-1}P_j}$ to the next direction vector $\overrightarrow{P_jP_{j+1}}$. The direction-change feature's degree (Fdc), is calculated by

$$Fdc = \frac{|D\theta|}{60} + 1 \tag{3}$$

where $D\theta$ ($-180° \leq D\theta \leq 180°$) is the angle of the direction change between $\overrightarrow{P_{j-1}P_j}$ and $\overrightarrow{P_jP_{j+1}}$.

Just as the 8-directional feature, the Fdc of each sampling point is also mapped to eight directions as Fig. 3(a). However, in direction-change feature, each Fdc is mapped to only one direction d_j^m from $D1, D2, \ldots, D8$. Concretely, it is the direction which the greater value between a_j^1 and a_j^2 corresponds to. Similarly, 8 direction-change pattern images $\{\dot{B}_d = [\dot{f}_d(x,y)], x, y = 1, 2, \ldots, 64, d = D1, \ldots D8\}$ are generated by setting $\dot{f}_{d_j^m}(x_j, y_j) = \max(a_j^1, a_j^2)$ and the remaining values as 0s. Then, the same computation is carried out on the generated direction-change pattern images to obtain the 512-dimensional direction-change feature vector.

We combine it with the 8-direction feature to form 1024-dimensional combined feature vector. Figure 4 shows 16 pattern images extracted from the Chinese character "Shi", where the first row represents the eight directional pattern images and the second row represents the direction-change pattern images.
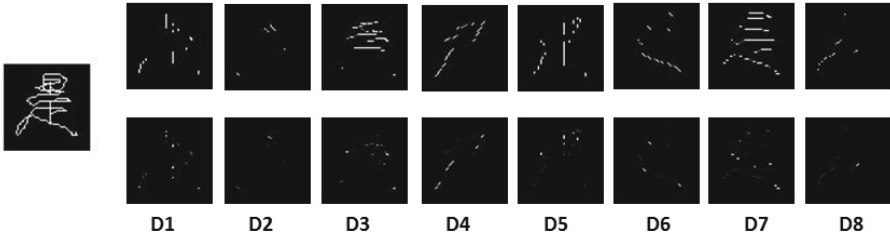

Fig. 4. Examples of pattern images for the directional and direction-change feature

4   Framework of Our Recognition System

We have implemented an recognition system for the in-air handwritten characters. Our system contains the following three stages of computations.

(1) Pre-processing. A series of pre-processing steps are used to reduce the noise and normalize the trajectory shapes of input characters to make the samples easier to recognize. First, we normalize the X-coordinates and Y-coordinates of the sampling points to a fixed size of 64 by 64 by linear mapping. Then, the coordinates of each sampling point are smoothed by computing the average of its neighbors, and we remove some redundant points to ensure that only one point left in the same position of the trajectory. Further, we exploit the dot density shape normalization method [16] to adjust the trajectory shape of the input Chinese character. Finally, the re-sampling step is carried out to generate a sequence of equidistance points.
(2) Feature Extraction. After the preprocessing step, we can extract the 1024-dimensional combined features. The related computational details have been presented in Subsects. 3.1 and 3.2.
(3) Two-Level Classifier. To make the classification more efficient, we exploit the Linear Discriminant Analysis (LDA) to learn a projection subspace so as to project the feature vector to a low-dimensional subspace. The projection axis learned by LDA helps to make prototypes more separable in the subspace. In LDA, we define within-class and between-class scatter matrices by S_W and S_B respectively and also define the optimal projection axis (discriminant vector) by $\mathbf{w}$. We then estimate $\mathbf{w}$ by maximizing the Fisher criterion:

$$J(\mathbf{w}) = tr((\mathbf{w}^T S_w \mathbf{w})^{-1}(\mathbf{w}^T S_B \mathbf{w})) \tag{4}$$

where $tr(\cdot)$ denotes the trace of matrix. This criterion considers the within-class and between-class scatter matrices and helps to make the data separable in the projected subspace. It can be shown that $\mathbf{w}$ is the solution to the generalized eigenvector problem $S_B w_i = \lambda_i S_W w_i, i = 1, 2, \dots$ where w_i denotes the eigenvector for the ith eigenvalue λ_i. By using LDA, the dimension of the feature space is reduced while different classes are separated. Also, the dimension reduction makes the following computation cost of training process decrease.

In our system, we design a two-level classifier to achieve both accuracy and efficiency. Our classifier is based on the Nearest Prototype Classifier (NPC) rule. For each unknown pattern, we label it by the class of the nearest prototype. The metric we use between samples and the prototypes is the the Euclidean distance. The first-level classier is the coarse classifier which aims to remove most impossible candidate classes with low computation cost. In the first level classifier, the combined feature vector is projected to 20-dimensional, and the nearest 450 prototypes are retained. Afterwards, in the second level classifier, we project the combined feature vector to 160 dimensional subspace, and then the distances are computed between the testing sample and prototypes retained by the first-level classifier. Finally, we sort these prototypes by the corresponding distances and generate the candidate label list.

5   Experimental Results

We evaluate the performance of the proposed feature on the IAHCC-UCAS2014 dataset which is constructed by ourself, since there is no related dataset publicly available. The dataset includes 3755 classes of Chinese characters, and each of them has 65 samples. The 3755 classes include all Chinese characters in GB2312-80 level-1 set, which makes our dataset challenging since works of other researchers [4,5] only cover limited number of classes in Chinese characters. Some of the samples in our dataset are shown in Fig. 5. As described in the previous section, it can be seen from Fig. 5 that the IAHCCR is technically difficult due to great variation of character's structure.

Fig. 5. Some in-air handwritten chinese characters from our IAHCC-UCAS2014 dataset

To evaluate the classification performance, the recognition accuracy on the testing data is of primary interest. We compare our recognition result with the other two features. The 8-directional feature refers to the method in [11] and the direction-change feature refers to the feature in [14]. The recognition accuracy is calculated by

$$R_k = N_k/N \tag{5}$$

where R_k denotes the top-k recognition accuracy of the system. For each testing sample, we generate a candidate label list for it by the two-level classifier. The top-k metric means we check candidate label list and find out if the right label is included in the top k label of the list. The N_k denotes the number of testing samples whose labels are properly included in the top-k candidate labels, while the N denotes the total number of the testing samples. We compare the accuracy of top 1, top 5, and top10, which are metrics widely used to compare Chinese character recognition performance. In our experiments, we randomly select 10 samples from each class as the testing samples and the remaining samples are used for training. The experimental results are summarized in Table 1.

It can be seen from the table that the proposed feature obtained better performance compared with the other two features according to top 1/5/10 metrics. It is also worth noting that the framework of our recognition system achieves 90.6 % accuracy on the top 1 metrics when dealing with the OHCCR

Table 1. Recognition accuracy comparison of three features on our in-air handwritten Chinese character dataset

Methods	Top1	Top5	Top10
direction-change	66.68 %	83.36 %	87.42 %
8-directional	68.81 %	84.80 %	88.52 %
Combined (Ours)	69.67 %	85.36 %	89.08 %

using dataset of SCUT-COUCH2009 [2]. The relatively low accuracy on the IAHCC-UCAS2014 dataset indicates that the IAHCCR is very challenging and it needs more research efforts in the future.

Compared with the 8-directional feature, the proposed method seems to bring more computational cost owing to combining the two features together. In practice, we apply the same pre-processing steps and direction vector extraction process to the two features in our recognition system so that the time cost can be reduced. We compare the time consumption (millisecond) during the feature extraction step and the recognition step. The experiments are performed on a desktop computer with 2.40 GHz CPU, and the recognition system is implemented using MATLAB. It can been seen from Table 2 that our combined feature results in little time consumption. Regarding the applied background of the IAHCCR, the increase of time consumption is negligible.

Table 2. Comparison of time consumption (millisecond) vs. 8-directional feature

Methods	Extraction	Recognition	Total
8-directional	96.5	10.7	107.2
Combined (Ours)	128.6	10.8	139.4

6   Conclusions

In this paper, we present a novel HCI interface for writing interaction. The writing behavior can be conducted in a 3D space and in a more natural and user-friendly way by using the Leap Motion Controller. We then propose a combined feature based on the 8-directional feature and the direction-change feature, and apply them to our in-air handwritten character recognition system. The performance of the combined feature is evaluated on our IAHCC-UCAS2014 dataset and the experimental results show that the combined feature can achieve better performance with reasonable computational cost.

References

1. Liu, C.-L., Jaeger, S., Nakagawa, M.: Online recognition of Chinese characters: the state-of-the-art. IEEE Trans Pattern Anal. Mach. Intell. **26**(2), 198–213 (2004)
2. Jin, L., Gao, Y., Liu, G., Li, Y., Ding, K.: SCUT-COUCH2009-a comprehensive online unconstrained Chinese handwriting database and benchmark evaluation. Int. J. Doc. Anal. Recogn. **14**(1), 53–64 (2011)
3. Liu, C.-L., Yin, F., Wang, D.-H., Wang, Q.-F.: Online and offline handwritten Chinese character recognition: benchmarking on new databases. Pattern Recognit. **46**(1), 155–162 (2013)
4. Feng, Z., Xu, S., Jin, L., Ye, Z., Yang, W.: Real-time fingertip tracking and detection using Kinect depth sensor for a new writing-in-the-air system. In: Proceedings of the 4th International Conference on Internet Multimedia Computing and Service, pp. 70–74 (2012)
5. Ye, Z., Zhang, X., Jin, L., Feng, Z., Shaojie, X.: A new writing experience: finger writing in the air using a kinect sensor. IEEE Trans. Multimed. **20**(4), 85–93 (2013)
6. Jin, X.-J., Wang, Q.-F., Liu, C.-L.: Visual gesture character string recognition by classification based segmentation with stroke deletion. In: Proceedings of 2nd IAPR Asian Conference on Pattern Recognition, pp. 120–124 (2013)
7. Hodson, H.: Leap motion hacks show potential of new gesture tech. New Sci. **218**(2911), 21 (2013)
8. Sutton, J.: Air painting with Corel Painter Freestyle and the leap motion controller: a revolutionary new way to paint. In: ACM SIGGRAPH 2013 Studio Talks, no. 21 (2013)
9. Hantrakul, L., Kaczmarek, K.: Implementations of the Leap Motion device in sound synthesis and interactive live performance. In: Proceedings of the 2014 International Workshop on Movement and Computing, pp. 142 (2014)
10. Khademi, M., Hondori, H.M., McKenzie, A.: Free-hand interaction with leap motion controller for stroke rehabilitation. In: CHI 14 Extended Abstracts on Human Factors in Computing Systems, pp. 1663–1668 (2014)
11. Bai, Z.-L., Huo, Q.: A study on the use of 8-directional features for online handwritten Chinese character recognition. In: Proceedings of the 8th International Conference on Document Analysis and Recognition, pp. 262–266 (2005)
12. Liu, C.-L., Zhou, X.-D.: Online Japanese character recognition using trajectory-based normalization and direction feature extraction. In: Proceedings of 10th International Workshop on Frontiers in Handwriting Recognition, pp. 217–222. Suvisoft (2006)
13. Ding, K., Deng, G., Jin, L.: An investigation of imaginary stroke technique for cursive online handwriting Chinese character recognition. In: Proceedings of the 10th ICDAR, pp. 531–535. Barcelona, Spain (2009)
14. Masayoshi, O., Yamamoto, K.: On-line handwriting character recognition using direction-change features that consider imaginary strokes. Pattern Recognit. **32**(7), 1115–1128 (1999)
15. Okamoto, M., Yamamoto, K.: On-line handwritten character recognition method using directional features and clockwise/counter-clockwise direction-change features. In: Proceedings of 5th International Conference on Document Analysis and Recognition, pp. 491–494 (1999)
16. Bai, Z.-L., Huo, Q.: A study of nonlinear shape normalization for online handwritten Chinese character recognition: dot density vs. line density equalization. In: Proceedings of 18th International Conference on Pattern Recognition, pp. 921–924 (2006)