2010 International Conference on Pattern Recognition

A New Rotation Feature for Single Tri-axial Accelerometer Based 3D Spatial Handwritten Digit Recognition

Yang Xue, Lianwen Jin*
Key Laboratory of Guangdong Wireless Communication and Terminals
School of Elec. & Info. Eng., South China University of Technology, China, 510640
**E-mail: lianwen.jin@gmail.com*

Abstract

A new rotation feature extracted from tri-axial acceleration signals for 3D spatial handwritten digit recognition is proposed. The feature can effectively express the clockwise and anti-clockwise direction changes of the users' movement while writing in a 3D space. Based on the rotation feature, an algorithm for 3D spatial handwritten digit recognition is presented. First, the rotation feature of the handwritten digit is extracted and coded. Then, the normalized edit distance between the digit and class model is computed. Finally, classification is performed using Support Vector Machine (SVM). The proposed approach outperforms time-domain features with a 22.12% accuracy improvement, peak-valley features with a 12.03% accuracy improvement, and FFT features with a 3.24% accuracy improvement, respectively. Experimental results show that the proposed approach is effective.

1. Introduction

In recent years, accelerometer-based human activity recognition has been carried out successfully [1, 2, 4, 7]. This is one of the most important research subjects in wearable computing and ubiquitous computing. Because it has important applications in the areas of health monitoring, context-awareness, and a new form of human-computer interaction, accelerometer-based human activity recognition has gained the attention of many research groups. At present, human activity recognition based on accelerometers includes activity status detection and gesture recognition. Activity status detection implies identifying certain daily activities, such as walking, running, standing, sitting and so on [1, 4]. Gesture recognition involves distinguishing certain simple phone gestures [5, 6], custom signs for controlling devices [2, 7, 9], and the digits from 0 to 9 [3, 8]. Because activity recognition can be formulated as a typical classification problem similar to many pattern recognition problems, feature extraction plays a crucial role during the recognition process. In general, most of the attempts to extract features from acceleration data can be classified into two categories, namely, time-domain features and frequency-domain features. Traditional widely used time-domain features include mean [1, 4, 7], variance or standard deviation [1, 4], correlation between axes [1, 7], peak-valley [8], and so on. The most popular frequency-domain features are FFT coefficients [1, 7].

Some researchers attach great importance to the recognition of handwritten digits from 0 to 9 in 3D space [3, 8], because 3D spatial handwriting has more advantages than 2D spatial handwriting. First, compared with a 2D handwriting input system, users can write freely without the limitation that writing should be on a specific plane such as a tablet. Second, besides the users, it does not require additional equipment, such as cameras. Third, it provides a novel input interface. Although much research on exploring gesture recognition based on a tri-axial accelerometer [2, 5-9] has already been reported in the literature, extraction of features from acceleration data remains a research challenge.

In this paper, we first propose a novel rotation feature from an acceleration signal for 3D spatial handwriting. Then, we present an algorithm based on the rotation feature to recognize 3D spatial handwritten digits. Finally, we perform the classification with an SVM and compare the performance of different features. The results show that the rotation feature can effectively recognize the 10 Arabic digits.

2. Feature extraction

Contrary to traditional handwriting input systems, we cannot extract the trajectory of 3D spatial handwriting. We can only obtain a 3D acceleration signal generated by the 3D accelerometer while writing in 3D space. By observing many 3D acceleration signals' waves, we found that a change in 3D acceleration direction causes the handwriting trajectory's clockwise or anti-clockwise rotation, the turning points of which are the corner points. As different writing movement changes occur while handwriting different characters, we expect to extract the feature that can reflect the movement changes in order to recognize 3D handwritten characters. In this paper, we refer to the features that effectively express the clockwise and anti-clockwise direction changes as rotation features.

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.1025

Let 3D accelerations obtained from a 3D accelerometer be $a_x(t)$, $a_y(t)$, and $a_z(t)$. According to a previous work [9], the number of anti-clockwise rotations is calculated as the vector product of the vectors at t and $t+1$ as shown in Eqs. (1) and (2). When the gesture motion is anti-clockwise from $A(t)$ to $A(t+1)$, 1 is added to P_g.

$$P_g = \sum_{t=1}^{S-1} u(|A(t) \times A(t+1)|) \tag{1}$$

$$u(n) = \begin{cases} 1, & n \geq 0 \\ 0, & n < 0 \end{cases}, \tag{2}$$

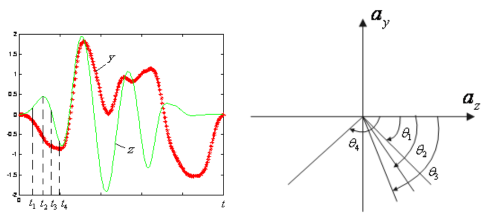
where $A(t) = [a_x(t), a_y(t), a_z(t)]$ is the tri-axial acceleration generated by a 3D accelerometer during handwriting. S is the length of $A(t)$.

In this study, the rotation feature is extracted from the following three two-dimensional vectors respectively:

$$\begin{aligned} A_{yz}(t) &= [a_y(t), a_z(t)] \\ A_{zx}(t) &= [a_z(t), a_x(t)] \\ A_{xy}(t) &= [a_x(t), a_y(t)] \end{aligned} \tag{3}$$

$A_{yz}(t)$, $A_{zx}(t)$, and $A_{xy}(t)$ are the projection vectors of the accelerations on the *y-z*, *z-x*, and *x-y* planes, respectively.

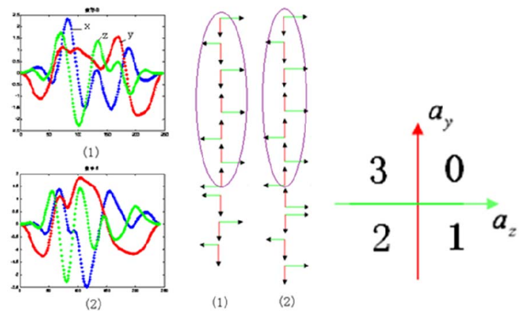
We use $A_{yz}(t)$ as an example to illustrate the extraction of rotation features. Fig. 1(a) shows the two-dimensional vector $A_{yz}(t) = [a_y(t), a_z(t)]$, which is the projection vector of acceleration generated by a 3D accelerometer during handwriting. The phases of $A_{yz}(t)$ in t_1, t_2, t_3, and t_4 are expressed as θ_1, θ_2, θ_3, and θ_4, respectively. The relation of the 4 phases is shown in Fig. 1(b). By observing Fig. 1(b), we found that rotation feature points can be selected according to the phase relation in the different moments. The phases of $A_{yz}(t)$ in t_i and t_{i-1} are expressed as θ_i and θ_{i-1}, respectively. Only when θ_i and θ_{i-1} are located in different quadrants, can $A_{yz}(t_i)$ be taken as a rotation feature point.



(a) (b)

Figure 1. (a) Acceleration in the y- and z-axes (b) Relations of the 4 phases

Fig. 2(a) shows the acceleration in the *x*-, *y*-, and *z*- axes generated by two users while handwriting the digit 8 in 3D space. Fig. 2(b) shows the rotation feature points $A_{yz}(t_i)$ extracted from the two-dimensional vector $A_{yz}(t)$, and then expressed as $[a_y(t_i), a_z(t_i)]$. In Fig. 2(b), the vertical axis represents the y-axis, the horizontal axis represents z-axis, while arrows point to the direction of the acceleration for each axis.



(a) (b) (c)

Figure 2. (a) 3D accelerations of 2 users (b) Rotation features in y-z plane (c) Coding the rotation features

As can be seen from Figs. 2(a) and (b), although there is a big difference between the 3D acceleration signals of the two users, the respective processes of acceleration changes in direction are the same (illustrated by the ovals). In this paper, the process of acceleration change in direction is extracted as the rotation feature for recognizing handwritten digits. To simplify numerical calculation, the rotation feature is coded as shown in Fig. 2(c).

The rotation features from the other two-dimensional vectors, $A_{zx}(t)$ and $A_{xy}(t)$, are extracted and coded in the same way. Thus, the rotation feature of a 3D acceleration is represented as

$$C = \{C_{yz}, C_{zx}, C_{xy}\} \tag{4}$$

where $C_i = c_1 c_2 \cdots c_n$, $C_i \in \{C_{yz}, C_{zx}, C_{xy}\}$, and $c_n \in \{0,1,2,3\}$. C_{yz}, C_{zx}, and C_{xy} are called the rotation feature codes of $A_{yz}(t)$, $A_{zx}(t)$ and $A_{xy}(t)$, respectively.

Since different samples have different length of rotation feature codes and the difference between them cannot be calculated directly, edit distance [10] is used to measure the difference between different rotation feature codes.

3. Recognition algorithm

Based on the rotation feature, an algorithm for 3D spatial handwritten digit recognition is given in Fig. 3.

We first reduce the noise using 1-D Gaussian smoothing and normalize the amplitude for each axis' data. According to Eq. (3), three two-dimensional vectors, $A_{yz}(t)$, $A_{zx}(t)$, and $A_{xy}(t)$ are obtained. The Rotation feature codes C_{yz}, C_{zx}, and C_{xy} are then extracted. Thus, we are able to extract tri-dimensional rotation feature codes from each sample. We randomly select k samples from the total samples as training samples. For the training samples, we place samples of the same class together. In each class, we calculate the sum of the normalized edit distances between each sample and the other samples. Then according to the sum of edit distance, we order the samples in increasing order, as the class model for each class. We calculate the edit distance between each sample and all class models, as the feature for each sample.

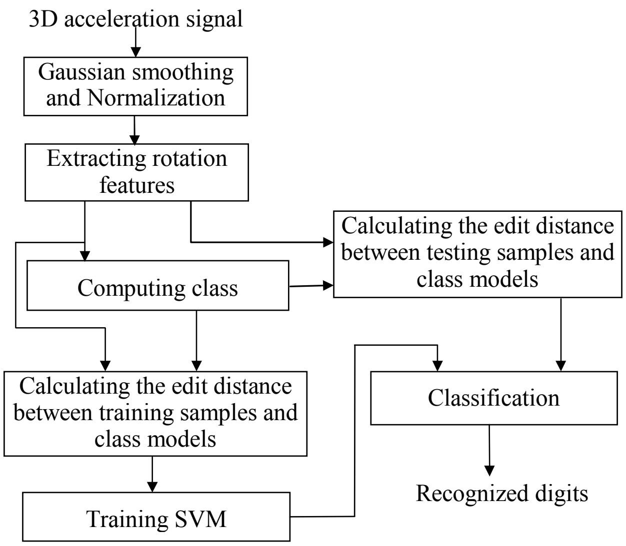


Figure 3. Algorithm framework

The classification algorithm used is a Support Vector Machine (SVM) [11] with One-versus-One strategy (OVO), in which a set of binary classifiers are constructed using corresponding data from two classes. While testing, we use the voting strategy of "Max-Wins" to produce the output.

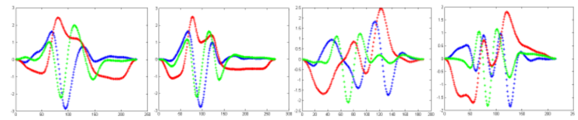
4. Experimental results and discussion

4.1. Data collection

We designed a data collection apparatus and collected data from 60 writers. Each writer freely wrote digits from 0 to 9 in 3D space three times. The total dataset contains 180 sets samples. In some previous studies, researchers strictly regulated the stroke order and angle of writing during data collection [3, 8]. However, we had no strict rules. Each writer was free to write according to his/her own writing habits. Fig. 4 shows an example of the acceleration signal in the *x*-, *y*-, and *z*- axes.

4.2. Performance comparison and analysis

To validate the rotation features (RF), we compared their performance against time-domain features (TF) [2], peak-valley features (PF) [8], and the widely used FFT coefficients [1, 7]. We extracted the first 64 FFT coefficients from each axis of acceleration data as FFT feature. We randomly selected 144 sets of the total dataset for training, with the remaining 36 sets for testing. The recognition results of three experiments are given in Table 1.



(a) "0" and "6" (b) "4" and "9"

Figure 4. Example of acceleration signal

Table 1. Accuracy based on four different features

Features	Accuracy of 3 experiments			Average
	1	2	3	
TF [2]	57.78%	58.89%	61.67%	59.45%
PF [8]	69.17%	69.44%	70%	69.54%
FFT [7]	77.78%	78.33%	78.89%	78.33%
RF	80.28%	81.94%	82.5%	81.57%

Table 1 shows that the proposed approach outperforms time-domain features with a 22.12% accuracy improvement, peak-valley features with a 12.03% accuracy improvement, and FFT features with a 3.24% accuracy improvement. The experiments thus confirm that the proposed approach is effective.

Table 2 shows the detailed recognition rates for 10 classes respectively. From these results, we can see that accuracy based on the rotation feature is higher for almost all class except class 7, whereas accuracy based on time-domain features or peak-valley features is considerably lower. The recognition accuracy using FFT features is slightly better than using rotation features. Integration of FFT coefficients and rotation features may produce even better performance, which merits our future study.

Table 2. Accuracy comparison of 10 classes

Class	Accuracy			
	TF	**PF**	**FFT**	**RF**
0	57.41%	60.19%	62.96%	73.15%
1	70.37%	82.41%	88.89%	85.16%
2	51.85%	64.82%	63.89%	74.07%
3	68.52%	69.44%	87.96%	96.3%
4	55.56%	70.37%	83.33%	81.48%
5	55.56%	72.22%	92.59%	92.59%
6	55.56%	70.37%	68.52%	77.78%
7	59.26%	72.22%	79.63%	76.85%
8	46.30%	65.74%	78.70%	81.48%
9	74.07%	67.59%	76.85%	76.85%
Average	**59.45%**	**69.54%**	**78.33%**	**81.57%**

To find which digits are relatively harder to recognize, we analyzed the confusion matrices. Table 3 shows the aggregate confusion matrix based on our new features for three experiments. It can be seen that digit "0" is often confused with digit "6", while digit "4" is often confused with digit "9". Because writers cannot get visual feedback of written trajectories during their writing, gestures with similar movement history, such as digits "0" and "6", "4" and "9", are frequently confused as shown in Fig. 4.

Table 3. Confusion matrix based on rotation feature

	Recognized as									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	79	3	1				22	1	2	
1	3	92	3			1	1	7		1
2		6	80	3		4	2	5	5	3
3			1	104		3				
4					88	2	2	2	3	11
5			2		3	100			3	
6	21						84	1	1	1
7	3	9	3	1	2			83	2	5
8	3		6	1	1	5	1	1	88	2
9	1	1	1		10		2	10		83

5. Conclusion

In this paper, we proposed a novel rotation feature for 3D spatial handwritten digit recognition extracted from tri-axial acceleration signals that can effectively express the clockwise and anti-clockwise direction changes of user's hand movement while writing in 3D space. A method for 3D spatial handwritten digit recognition based on the rotation feature is presented. First, we extracted the rotation feature of the handwritten digit and coded the feature. Then, we computed the class model for each class using training samples. Thereafter, we computed a normalized edit distance between the digit and class model as the feature for the digit. Finally, we performed classification with an SVM classifier. The average accuracy of 10 digits using the proposed rotation feature is 81.57%, which is better than using time-domain features, peak-valley features, and FFT features. The experimental results confirm the effectiveness of the proposed approach.

Acknowledgment

The authors would like to thank the data collecting volunteers from SCUT.

References

[1] L. Bao and S. S. Intille. Activity Recognition from User-Annotated Acceleration Data. The 2nd International Conference on Pervasive Computing, 2004.

[2] S. Thomas, P. Benjamin, et al. Gesture Recognition with a Wii Controller. Proc. of the 2nd Int. Conf. on Tangible and Embedded Interaction, pp.11-14, February 2008.

[3] E.-S. Choi, W.-C. Bang, et al. Beatbox Music Phone: Gesture Interactive Cell phone using Tri-axis Accelerometer. IEEE Int. Conf. on Industrial Technology, pp.97-102, December 2005.

[4] M. Ermers, J. Pärkkä, et al. Detection of Daily Activities and Sports With Wearable Sensors in Controlled and Uncontrolled Conditions. IEEE Transactions on Information Technology in Biomedicine, vol.12, pp.20-26, January 2008.

[5] V-M. Mäntylä, J. Mäntyjärvi, et al. Hand Gesture Recognition of a Mobile Device User. IEEE Int. Conf. on Multimedia and Expo, pp.281-284, 2000.

[6] Zhenyu He, Lianwen Jin, et al. Gesture Recognition based on 3D Accelerometer for Cell Phones. IEEE Asia Pacific Conf. on Circuits and Systems, pp.217-220, 2008.

[7] Wu Jiahui, Pan Gang, Zhang Daqing, et al. Gesture Recognition with a 3-D Accelerometer. Ubiquitous Intelligence and Computing, vol.5585, pp. 25-38, 2009.

[8] S.-J. Cho, E.-S. Choi, et al. Two-stage Recognition of Raw Acceleration Signals for 3-D Gesture-Understanding Cell Phones. The 10th Int. Workshop on Frontiers in Handwriting Recognition, 2006.

[9] H. Sawada, S. Hashimoto. Gesture Recognition Using an Acceleration Sensor and Its Application to Musical Performance Control. Electronics and Communications in Japan, vol.80, pp.9-17, 1997.

[10] V.I. Levenshtein. Binary Codes Capable of Correcting Deletions, Insertions, and Reversals. Soviet Physics Doklady, vol.10(8), pp.707-710, 1966.

[11] V. Vapnik. The Nature of Statistical Learning Theory, Springer Press, New York, 1999.