Key-press Gestures Recognition and Interaction Based on SEMG Signals

Juan Cheng Xiang Chen Zhiyuan Lu
Dept. of Electronic Sci. & Technol.
University of Science & Technology of China, 230027, China
xch@ustc.edu.cn
jncheng@mail.ustc.edu.cn
luzhiy@mail.ustc.edu.cn

Kongqiao Wang
Nokia Research Center,
NOKIA (CHINA) Investment CO. LTD.
Beijing, 100080 China
Kongqiao.Wang@nokia.com

Minfen Shen
School of Engineering
Shantou University, 515063, China
mfshen@stu.edu.cn

ABSTRACT
This article conducted research on the pattern recognition of key-press finger gestures based on surface electromyographic (SEMG) signals and the feasibility of key-press gestures for interaction application. Two sort of recognition experiments were designed firstly to explore the feasibility and repeatability of the SEMG-based classification of 16 key-press finger gestures relating to right hand and 4 control gestures, and the key-press gestures were defined referring to the standard PC keyboard. Based on the experimental results, 10 quite well recognized key-press gestures were selected as numeric input keys of a simulated phone, and the 4 control gestures were mapped to 4 control keys. Then two types of use tests, namely volume setting and SMS sending were conducted to survey the gesture-base interaction performance and user's attitude to this technique, and the test results showed that users could accept this novel input strategy with fresh experience.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces–*Input devices and strategies, Interaction styles.*

General Terms
Verification

Keywords
Human computer interaction, key-press finger gesture, virtual keyboard, electromyographic.

1. INTRODUCTION
Nowadays, PDAs, smart mobile phones and wearable computers become more popular, due to the portability, the size of these devices are shrinking, which brings challenges to traditional input facilities. Firstly, those traditional input facilities are difficult to carry around. Secondly, users' hands are sometimes occupied and cannot operate. Hence, speech, handwriting, and gesture based input manners draw more attention [1~3].

Gesture-based input style is of great importance to be researched, especially for particular occasion. For instance, someone is in the kitchen, with his/her hands not available when a phone call is ringing. According to the traditional manner, he/she will miss the phone or has to stop to pick up the phone, whereas gesture SEMG interaction will avoid above problems.

Currently, many sensing technology, such as glove-based, body-suit-based, and external cameras and computer vision technique [4] are implemented to capture gestures. SEMG sensors do have unique superiority, especially in capturing subtle gestures. That's because SEMG signals are superposition of motor unit action potentials (MUAP) from myoelectrical contraction on physical surface [8], and contain vivid information of gesture execution [5].

Many institutions research on recognition gesture SEMG related to wrist, elbow and finger action. For example, Kevin R. Wheeler et al. [6] executed research on recognition of forearm gesture SEMG. Meanwhile, they accomplished interaction of simulated aircraft by virtual joystick, as well as virtual numeric keypad, with numbers 0 to 9 and an "Enter". T.Scott Sapona et al. [7] produced a system that classifies four-finger gestures, with accuracies varying from 79% to 88%.

Our research focuses on simulated phone interaction system with virtual keyboard based on key-press finger gestures. Two kinds of experiments are firstly designed to testify the feasibility of pattern recognition of key-press finger gestures. Then, use tests are implemented to survey gesture SEMG interaction performance and users' attitude to this technique.

In the following sections, methods of gesture recognition are displayed, and a real-time platform of simulated phone will be proposed to evaluate gesture SEMG interaction performance. Then pattern recognition and analysis is demonstrated, and use tests will follow. Ultimately, a conclusion is given.

2. APPROACH
2.1 Gesture Definition & Electrode Placement
In this paper, referring to standard PC keyboard, 16 sorts of key-press finger gestures, related to five fingers of right hand, are

defined. As shown in Figure 1, keys of identical color are pressed by same finger, as purple keys (Y, U, H, J, N, M) are controlled by index finger. Considering interaction with simulated phone, 4 extra control gestures, described in Table 1, are included. Totally 20 gestures are named with 2 letters in Table 2. Each letter key is named with two identical letters, as letter Y is named with YY. When referring to symbol keys, the first and last letters of spelling are taken to represent, like period (.) is named with PD. Besides, relaxation is defined, totally relaxed, as illustrated in Figure 1.

All key-press finger gestures are executed on a paper keyboard similar to actual PC keyboard. Gestures begin and end up with relaxation state. When a key-press gesture is performed, only the finger, which presses, strikes hard.

Figure 1. Illustration of right hand initial placement

According to the above gesture definition, the related muscles are mainly *M. flexor carpi radialis*, *M. flexor carpi ulnaris, M. palmaris longus* and *M. extensor carpi radialis longus.* [8] Therefore, 6 channel electrodes are placed upon these positions. The electrode placement refers to Figure 2. Signals captured by electrodes are converted to digital data by NIDAQ-6010 PCMCIA acquisition card，and the sampling rate is 1000Hz.

Table 1. Control Gesture Definition

Control Gesture	Description	Control Key
HG	**H**and **G**rasp	OK
PE	**P**alm **E**xtension	Return
WE	**W**rist **E**xtension	Scroll Up
WF	**W**rist **F**lexion	Scroll Down

Figure 2. Scheme of SEMG Electrodes Placement

2.2 Gesture SEMG Recognition Algorithm

All experiments are executed on real-time platform. Gesture recognition process of SEMG signal includes detection of gesture action segmentation, feature extraction, and classifier design.

2.2.1 Detection of gesture action segmentation

In our research, differential moving average algorithm is applied. The differential values of each channel SEMG signal are obtained referring to formula (1). And, moving average algorithm [9] is used to identify the beginning and ending point of gesture action.

$$S_{averaged}(i) = \frac{1}{C}\sum_{k=1}^{C}[S_k(i+1) - S_k(i)]^2 \tag{1}$$

2.2.2 Feature extraction

Feature extraction is to effectively represent differences between gestures, using fewer feature vectors. For present, SEMG signal feature parameters are extracted from time domain, frequency domain and time-frequency domain [10].They have a positive effect on producing good recognition performance. Considering accuracy and real-time property, MAV and third-order AR coefficients for each channel were extracted [11].

2.2.3 Classifier design

Many classifiers are designed to recognize gesture patterns [12], Classic linear classifiers are simple, stable and good performance of real-time system realization, while non-linear classifiers, such as Artificial Neural Network (ANN), Hidden Markov Model [HMM], take advantage of high classification rates, but large computation. Here, Linear Discriminant Classifier (LDC) [13] adopted to accomplish the real-time recognition of gestures.

3. PATTERN RECOGNITION & ANALYSIS

Six healthy subjects (three female), named after four letters (See Table 3), with ages varying from 22 to 25 years, were voluntarily involved in experiments. During the process of data acquisition, all the subjects sat comfortably on chairs, with their hands put on table in relaxation. Afterwards, 20 gestures, repeated about 20 times, were executed in sequence in a natural way. Each subjects finished 6 times experiments in different day in a month.

Firstly, SEMG signals vary between individual. Secondly, even the same individual, SEMG signals are different for physiological and psychological variations, as well as electrode displacement. Thereby, before interaction, we designed two types of data processing and analysis experiments to explore the feasibility and repeatability of the SEMG-based finger gesture recognition, and to select control gestures for the later interaction application.

The feasibility of key-press gesture recognition is explored by conducting same-time experiment with the training and testing data from the same day. Half of the data samples are used to train the classifier, and the rest for testing. Table 2 give classification rates averaged on the six subjects of the same-time experiment. As Table 2 shows, the average classification rate reaches 87.6% with 11.1% standard deviation (std.) and 15 (marked with boldface) out of 20 rates are above 84%, which demonstrate the defined key-press finger gestures and control gestures are mostly well recognized. However, because of the muscles which control the movement of index finger and middle finger is so close that gestures related to the two fingers are not perfectly discriminated. For instance, UU and II are usually classified mistakenly for each other. Meanwhile, the gestures, which are controlled by the same finger, are sometimes confused for each other. For example, NN and MM are prone to be assorted into each other falsely.

The repeatability of key-press gestures recognition is explored by conducting cross-time experiments with training and testing data from different time section. Data from the former *n* days are used for training, while those from the $n+1^{th}$ day for testing. Five days experiments had been implemented, so *n*=1~5. Table 3 shows the classification rates of same-time and cross-time experiments of each subject. From Table 3, with the increment of training samples, classification rates are obviously improved. When the cross time is 5, the classification rates are close to or even better than that of the same-time condition.

Table 2. Averaged key-press gestures classification results for the same-time experiment

	YY	UU	II	OO	PP	HH	JJ	KK	LL	SN	NN	MM	CA	PD	SH	SP	HG	PE	WE	WF	Avg
Mean	**94.3**	64.1	**88.1**	**85.1**	**94.2**	**93.2**	**84.1**	**84.0**	**85.8**	**97.0**	82.1	64.6	75.7	73.7	**92.3**	**93.1**	**100.**	**100.**	**100.**	**100.**	**87.6**
Std.	**6.6**	18.5	**9.1**	**12.5**	**8.4**	**4.4**	**11.0**	**10.7**	**10.5**	**3.5**	10.3	4.3	17.2	7.1	**7.9**	**9.4**	**0.0**	**0.0**	**0.0**	**0.0**	**11.1**

Table 3. Classification rates of the same-time & cross-time experiments

	CHJN	WDXG	TJXN	DHAO	WAXH	MACP	Avg. (%)
Same-time	91.7 ± 13.4	85.7 ± 13.3	88.0 ± 15.8	85.6 ± 16.1	87.5 ± 12.8	85.1 ± 12.8	87.6 ± 11.1
Cross-time n=1	84.4 ± 24.6	67.4 ± 40.0	76.8 ± 30.4	76.9 ± 30.5	83.0 ± 24.5	78.6 ± 20.6	77.8 ± 6.0
Cross-time *n*=2	89.2 ± 21.4	88.5 ± 18.7	77.9 ± 35.0	80.8 ± 32.8	81.2 ± 23.4	82.0 ± 19.4	83.3 ± 4.6
Cross-time *n*=3	84.0 ± 30.7	87.8 ± 17.3	83.6 ± 27.8	80.3 ± 24.3	87.2 ± 17.5	83.0 ± 20.5	84.3 ± 2.8
Cross-time *n*=4	91.8 ± 21.4	88.0 ± 17.8	88.3 ± 17.8	90.6 ± 18.4	89.4 ± 15.7	84.8 ± 20.5	89.2 ± 2.0
Cross-time *n*=5	94.5 ± 14.6	88.8 ± 13.2	88.5 ± 22.0	91.3 ± 18.3	92.1 ± 10.9	86.3 ± 14.1	89.9 ± 3.3

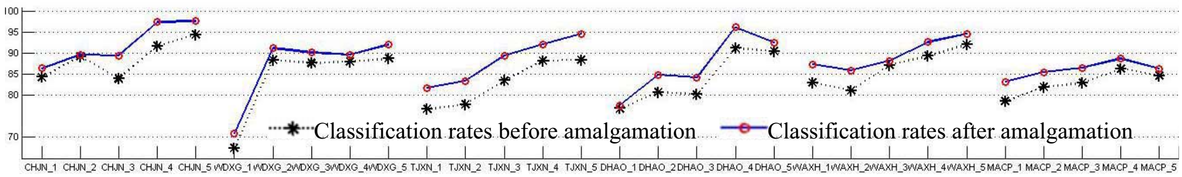

Figure 3. The average classification rates of gesture recognition before and after amalgamation

To further improve the classification performance, we adopted an amalgamation hypothesis. When gesture UU is misidentified to II, or II for UU, we artificially consider the classifications are both right. So do for NN and MM gestures. Under this hypothesis, the average classification rates are increased, shown in Figure 3. We also find the rates increased with cross-time number increment. The phenomenon can be explained as follows. 1) Although the classification rates decrease with electrode displacement (See Table 3, when cross-time n=1, the rates are the lowest), with the increment of cross-time number, samples cover most of the electrode displacements, the trained classifier becomes no more sensitive to electrode displacement. 2) All the experiments were executed on real-time platform, and results were simultaneously shown to subjects, which help subjects execute proper gestures. This equals to a training process. When subjects become familiar enough with the gesture execution, the results will keep steady.

From Table 2, Table 3 and Figure 3, we can discover that, despite of the differentia of subjects' physiological and psychological situation, as well as electrode displacement, the results positively proved the repeatability of the gesture SEMG recognition. Consequently, the key-press finger gestures, which keep the good performance both in the same-time and cross-time experiments, were selected as numeric input keys, and the control gestures were picked up as control keys in the interaction system.

4. USE TESTS

Our research target is to implement gesture-based interaction with actual mobile phones, before that, a simulated phone interaction platform is established to verify performance of this novel input strategy. Volume settings and SMS sending were designed as use test tasks. Based on the experimental results in last section, 10 quite well recognized gestures are selected as numeric input keys, and the mapping of the selected gestures and numeric input keys is illustrated in Figure 4 (marked with boldface). Four control gestures are mapped to control keys as displayed in Figure 4 (marked with normal font).

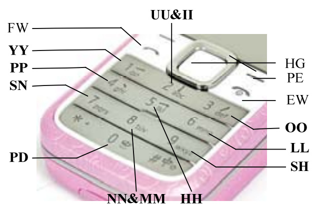

Figure 4. Numeric input keys mapping with key-press finger & control keys mapping with control gestures

The volume setting task is to accomplish volume setting using merely 4 control gestures. The simulated phone operation is as follows: Enter the main menu, scroll up or down to select volume setting, enter and increase the volume, finally return. The related control gestures are: HG for enter, EW for scroll up, FW for scroll down and PE for return. Meanwhile the SMS sending task is to send a short message to a certain number, using both control gestures and numeric gestures. When referring to control keys, corresponding control gestures are executed, and phone number and message context are edited by above selected key-press gestures. The supposed gesture execution numbers of first task is 10, and 19 for second task, including 13 key-press gestures and 6 control gestures.

Six same subjects involved in user tests. In both test experiments, some parameters are recorded to evaluate performance of the simulated mobile phone interaction system. The parameters mainly include: 1) How many gestures are supposed to be executed to accomplish the task? 2) How many gestures are indeed executed to accomplish the task? 3) How much time is consumed to fulfill the task? And 4) Is the subject accustomed to this interaction method? The parameters are recorded three times, and the averaged results are shown in Table 4.

Table 4: Use test parameters recording

	Volume Setting		SMS Sending	
	Indeed Num	Avg. Time(s)	Indeed Num	Avg. Time(s)
CHJN	10	11.76	21	27.26
WDXG	10	14.71	22.3	28.27
TJXN	10	13.42	24	31.56
DHAO	10	13.71	26.3	37.43
WAXH	10	12.64	23.7	31.49
MACP	10	14.33	25.7	36.00
Avg.	10	13.43	23.8	32.00

From Table 4, because of the 100% classification rates of control gestures, subjects have no difficulty in first task. The indeed gesture execution numbers is equal to that of supposed ones (10). Furthermore, subjects were interested in the interaction system.

For the second task, subjects found it slightly troublesome. Firstly, not all the key-press finger gestures are perfectly distinguished that it takes much more time to correct errors. Second, during interval of gesture execution, some unexpected gestures showed up, causing abrupt operation of interaction. However, despite of this, subjects were fond of this interaction, which brought fresh experience to daily life and convenience to hand-busy occasion.

In addition, during experiments on interaction system, we found that, the paper keyboard is not necessarily needed. Subjects can conduct gestures fluently and correctly even without paper, which means an invisible virtual keyboard can be carried by at any time.

5. CONCLUSION

This paper conducted research on pattern recognition of key-press finger gestures and control gestures based on SEMG signals, and the feasibility of gesture-based interaction with simulated phone. Two categories of recognition experiments were designed to verify the feasibility and repeatability of SEMG-based gesture recognition. The results showed that key-press finger and control gestures were well-recognized and prone to be used in interaction. Moreover, two types of use tests were implemented, and the results showed that users could accept this novel input strategy.

6. ACKNOWLEDGMENTS

Our research is supported by the Rearing Fund for Innovative Research Team of Univ. of Sci. & Technol. of China, by National High Technol. of Research and Development Program of China (863 Program) under Grant 2009AA01Z322, and supported by National Nature Sci. Foundation of China (NSFC) under Grant 60703069.

We would like to thank volunteers for their help in the experiments.

7. REFERENCES

[1] Deng, L., Wang, Y., and Wang, K. 2004. Speech and language processing for multimodal human-computer interaction. *J. VLSI Signal Processing Syst. Signal Image Video Technol.* 2-3(Feb-Mar, 2004) 161-187

[2] Alejandro, J. and Nicu, S. 2007. Multimodal human-computer interaction : A survey. *J. Comput. Vision Image Understanding.* 1-2(Oct, 2007) 116-134

[3] Naik, G. R. Kumar, D. K. and Arjunan S. P. 2008. Multi modal gesture identification for HCI using surface EMG. In *Proceedings of the 12th International Conference on Entertainment and Media in the Ubiquitous Era* (Tampere, Finland, October 07 - 09, 2008). ACM New York, NY, USA, 90-94

[4] Mitra S. and Acharya T. 2007. Gesture recognition: a survey. *IEEE Trans. Syst. Man Cybernetics*. 37, 3, (May, 2007) 314-324

[5] Costanza, E. Inverso, S. A. and Allen R. Toward subtle intimate interfaces for mobile devices using an EMG controller. 2005. In *Proceeding of the SIGCHI Conference on Human Factors in Computing Systems*.(Portland, Oregon, USA April 02 - 07, 2005) CHI'05. ACM New York, NY, USA, 481-489

[6] Wheeler, K. R. and Jorgensen, C. C. 2003. Gestures as input: Neuroelectric Joysticks and Keyboards. *IEEE Pervas. Comput*. 2, 2, (Apr, 2003) 56-61

[7] Saponas, T. S. Tan, D. S. and Morris, D. 2008. Demonstrating the feasibility of using forearm electromyography for muscle-computer interfaces. In *Proceeding of 26th annual SIGCHI Conference on Human Factors in Computing Systems*. (Florence, Italy, April 08 - 15, 2008) CHI'05. ACM New York, NY, USA, 515-524

[8] Konrad, P. 2005. The ABC of EMG: A practical introduction to kinesiological electromyography. Version 1.0, Noraxon. Inc. USA

[9] Zhang, X. Chen, X. and Zhao, Z. Y. 2007. Research on gesture definition and electrode placement in pattern recognition of hand gesture action SEMG. In Proceeding of 1st International Conference of Medical Biometrics. (Hong Kong, China, Jan 04 - 05, 2007)

[10] Oskoei, M. A. and Hu, H. S. Myoelectric control systems: a survey. 2007. *J. Bionedical Signal Processing Control.* 2, 4, (Oct,. 2007), 275-294

[11] Konx, R. and Brooks, D.H. Classification of multifunction surface EMG using advanced AR model representations. 1994. In Proceedings of 20th Annual Northeast.(Boston, MA, Mar 17 - 18, 1994)

[12] Ahsan, M. R. Ibrahimy, M. and Khalifa, O. O. 2009. EMG signal classification for human computer interaction: a review. *J. Scientific Research.* 33, 3 (Jul, 2009), 480-501

[13] Liu, C. J. and Wechsler, H. 2000. Robust coding schemes for indexing and retrieval from large face databases. *IEEE Trans. Image Processing.* 9, 1, (Jan, 2000), 132-137