

# A Japanese Input Method for Mobile Terminals Using Surface EMG Signals

Akira Hatano<sup>1</sup>, Kenji Araki<sup>1</sup>, and Masafumi Matsuhara<sup>2</sup>

<sup>1</sup> Graduate School of Information Science and Technology,  
Hokkaido University, Kita-ku Kita 14 Nishi 9, Sapporo, 060-0814, Japan  
`{a_hatano, araki}@media.eng.hokudai.ac.jp`

<sup>2</sup> Department of Software and Information Science,  
Iwate Prefectural University, Iwate, 020-0193, Japan  
`masafumi@soft.iwate-pu.ac.jp`

**Abstract.** The common use of mobile terminals is for text input. However, mobile terminals cannot be equipped with sufficient amount of keys because of the physical restrictions. To solve this problem we developed an input method using surface electromyogram (sEMG), treating arm muscle movements as input signals. This method involves no physical keys and can be used to input Japanese texts. In our experiments, the system was capable of inputting Japanese characters with a finger motion recognition rate of approximately 80%.

**Keywords:** input method, surface electromyogram, new generation interfaces, human interface.

## 1 Introduction

In recent years, mobile terminals have been rapidly miniaturized and improved in performance, and they are often used to input texts. However, such terminals cannot be equipped with sufficient amount of keys because of their physical limitations. Therefore, it is difficult to speed up the input process. There are several studies on effective input methods using only few keys aiming at solving this problem [1,2,3], but the problem of the need for physical keys still remains. Using audio input as a substitute for a physical terminal eliminates the need for physical keys, but introduces new problems, such as recognition accuracy in loud environments and privacy issues.

In this paper, we describe our development of an input method using surface electromyogram (sEMG) which allows treating arm muscle movements as input signals. This method uses no physical keys and enables the input of Japanese characters. EMG signals are electrical signals from a muscle, and are associated with muscle activity. The sEMG signals are measured by receiving signals from a sensor placed on the skin. There have been numerous related studies on muscle fatigue measurement [4], prosthetic hand control [5] and human machine interfaces [6,7,8]. The sEMG signals of hand muscle activities are measured from an

area between the wrist and the lower arm, and the user can use the interface without any physical restrictions.

For these reasons, the sEMG should be well-suited as a wearable computer input method. We have been developing a system that recognizes user arm muscle motions using the sEMG signals and translates them to text input [9]. In this paper, we describe a system that recognizes user hand motions using sEMG signals, recognizes 6 different types of input based on those signals and then converts the input into Japanese text in real time. We also introduce the results of an experiment evaluating the number of key presses required to input evaluation texts. Figure 1 shows a screenshot of the system interface.

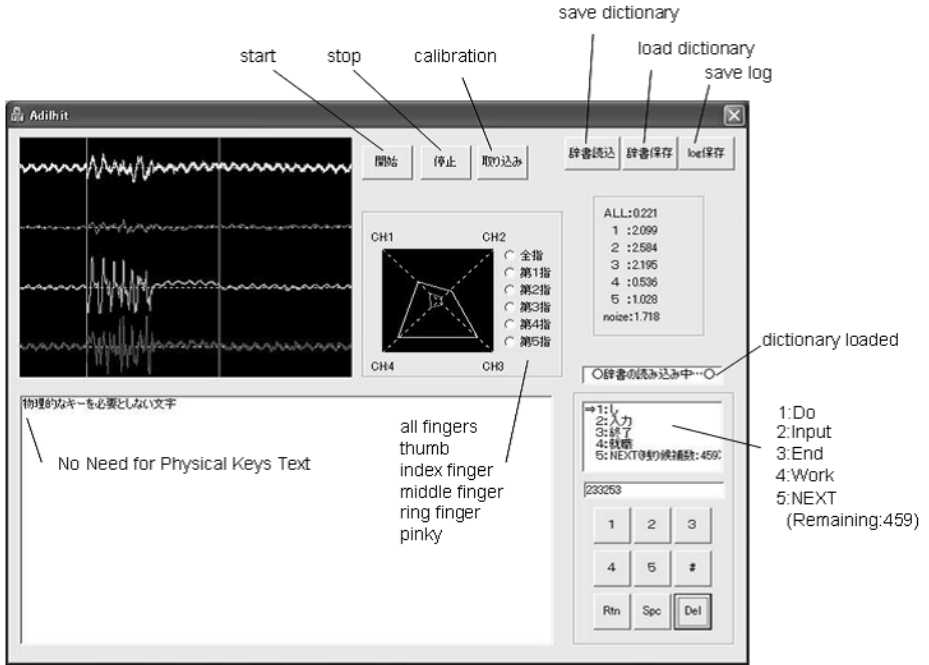
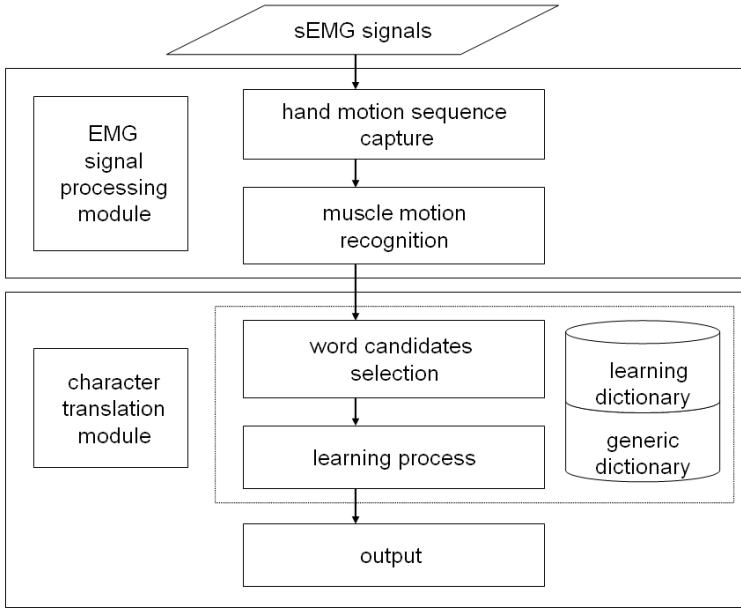


Fig. 1. Screenshot of the system interface

## 2 System Outline

Figure 2 shows an outline of the experiment system. The sEMG signal processing module uses sensors on the user's arm to capture sequences of arm muscle motions and recognizes finger motions. It then converts the signals to numerical pseudo keys and sends them to the character translation module which shows word candidates corresponding to the input. Selected words are then output in the text box.

The character translation process uses a generic dictionary and a learning dictionary for choosing word candidates. The generic dictionary contains the



**Fig. 2.** Outline of the system

10,000 most frequent words from the "Web Japanese N-gram Version 1 [10]". The learning dictionary registers words from the generic dictionary when they are selected by the user. It also keeps information on input frequency, surrounding word n-gram, etc. As it is being updated after every input word, the system can display word candidates adapted to particular users.

### 3 The EMG Signal Processing Module

In this section we describe the process through which the system recognizes finger flexion using the sEMG signals. In our experiment an sEMG sensor is installed on the flexor digitorum superficialis muscle area and the flexor pollicis longus muscle area that relate to finger flexion as shown in Figure 3. The device for measuring sEMG used in our study is a Personal-EMG [11]. The EMG signal processing module uses four input channels to recognize which finger has been flexed.

#### 3.1 Capturing Hand Motion Sequences

Figure 4 shows the raw sEMG signal difference between a flexed finger and an extended finger. However, the values of the raw sEMG signal oscillate heavily which makes it difficult to tell sequential finger movements apart. To compensate for the oscillations, the system integrates the signal from each of the four channels using the formula shown in (1).

$$S_{x(ch)} = \int_{-0.1(sec)}^0 |e(t + \tau)| d\tau \quad (1)$$

The result of calculating the integral of the sEMG signal is shown in Figure 5. The system captures sequences of hand motions based on intervals of motion/no-motion. If the integrated value in any channel is above a threshold value, the system determines that the user is moving his fingers. Similarly, if the integral value drops below the threshold value, the system determines that the user has stopped moving his fingers. Figure 6 shows an example of a captured hand motion sequence. The system recognizes the interval between the two vertical lines as one motion sequence.

### 3.2 Recognizing Hand Motions

Once the hand motion sequence has been captured, as described in 3.1, the system determines which finger (or combination of fingers) has been flexed. During its initialization steps, our system collects and stores calibration data for each user. The user is asked to flex each finger one at a time and all fingers at once. The calibration data consists of values for finger flexion for each finger (and all fingers at once) on all four channels, as shown in Table 1. To determine which finger has been flexed, the system first calculates the differences between the two most extreme values from each channel. It then calculates the Manhattan distance between the calculated differences and the flexion values in the calibration data and assigns a hand motion based on the smallest distance. Figure 7 shows the difference values calculated for the signal captured in Figure 6 and Figure 8 shows the calculated Manhattan distance for each hand motion. In this example, the system classifies the hand motion as an all-finger flexion, as it has the smallest Manhattan distance.

## 4 The Character Translation Module

### 4.1 Allocation of Japanese Syllables

As before, the system recognizes six types of motions. When inputting characters, finger motions are assigned to pseudo keys. The system assigns thumb flexion

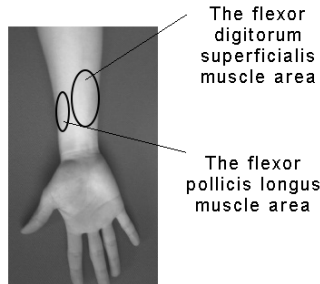
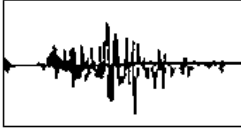
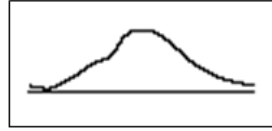


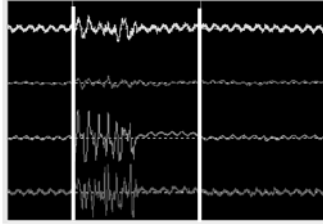
Fig. 3. Measurement points



**Fig. 4.** Raw sEMG signal



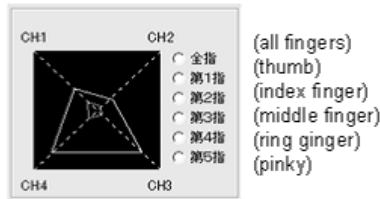
**Fig. 5.** Integrated sEMG signal



**Fig. 6.** An example of a captured hand motion sequence

**Table 1.** Calibration data in experiment

flexion	CH1	CH2	CH3	CH4
Thumb	4.983	4.219	1.545	0.818
Index finger	1.772	2.546	4.878	2.473
Middle finger	1.318	4.299	4.426	1.067
Ring finger	1.208	1.689	4.478	1.768
Pinky	1.248	3.562	1.360	2.898
all fingers	4.470	4.001	3.870	4.236



**Fig. 7.** Amplitude values for each channel

to "key 1", index finger flexion to "key 2", middle finger flexion to "key 3", ring finger flexion to "key 4", pinky flexion to "key 5", and flexion of all fingers to "enter key". The "enter key" is used to display and choose from available word candidates. "Keys" 1 through 5 are used to input Japanese syllables. Japanese syllables consist of one out of five possible vowel sounds and one out of ten possible consonant sounds. Since both the number of vowels in Japanese and the number of input keys are five, each vowel is assigned to a specific key (for example, "key 1" is assigned to vowel "a"). Thus, each key has many syllables assigned to it (for example, "key 1" corresponds to the syllables "a", "ka", "sa",



Fig. 8. Differences between current sequence and model calibration data

”ta”, etc.) and the number of times the user has to activate an input key to select one syllable is large. In order to achieve efficient input we use the Number-Kanji Translation method [12], [13]. In this method the user selects a syllable set with a single key press (for example, the syllable set corresponding to ”key 1”) and disambiguation is left for later. On one hand, the number of key presses is much smaller than that of a general method, such as the letter cycling input used in Japanese cell phones. On the other hand, the number of word candidates gets larger, and displaying a correct word candidate becomes a problem. To solve this problem, the system utilizes two dictionaries, and uses them effectively.

4.2 The Generic and Learning Dictionaries

The generic dictionary contains 10,000 words. These are sorted in descending order of frequency. Word candidates are displayed first in order of the longest match and then in dictionary order (by frequency). When there are word candidates available, the system displays up to four word candidates at a time as shown in Figure 9. The user selects a word candidate using ”key 1-4” and accepts the candidate with the ”enter key”. To display the next 4 candidates the user selects ”key 5” followed by the ”enter key”. Inputting ”key 5” and ”enter key”, the next four word candidates are displayed.

Words selected by the user are removed from the generic dictionary and are saved in the learning dictionary. Words contained in the learning dictionary are preferred over other words when displaying word candidates. In order to sort the words in the learning dictionary we use a Credibility Evaluation Function similar to [9]. Our version of this function is shown in (2).



Fig. 9. The displayed word candidates

$$CEF = \alpha \times ND + \beta \times UN + \gamma \times LN + \delta \times RD \quad (2)$$

There are three changes compared to the method described in [9]. First, we have removed class structures, so all words are now candidates. Second, in order to adopt a method that selects words, two measures were removed from CEF - Frequencies of Correct Translation and Frequencies of Erroneous Translation. Third, the measure Word Uni-gram Frequency (UN), Length of Word Candidate (LN) and Degree of Recency (RD) were added to CEF.  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$  are coefficients. The Degree of Neighboring Character Strings (ND) is calculated from the user's input history of word bi-grams, tri-grams, etc.

## 5 Evaluation Experiment

In our experiment we evaluate the hand motion recognition rate, the number of key presses and the average rank of the correct word candidate. The documents used in this experiment are author's mobile phone e-mails.

All words in these documents already existed in the the generic dictionary. For evaluating the recognition rate of hand motions and the number of key presses we used 72 e-mail messages (10.3 Japanese characters on average), and for the evaluation of the average rank of the correct word candidate we used 144 e-mail messages. In its initial condition the system's learning dictionary is empty and the coefficients of the CEF are set to  $\alpha=10, \beta=0.1, \gamma=10, \delta=0.1$  based on previous experimental results.

In the evaluation of the hand motion recognition rate, if the user's finger flexion is the same as what the system recognizes, it counts as a correct recognition, otherwise it counts as a false recognition. In this evaluation, we measure the recognition rate of hand motions only in the Japanese syllables selecting experiment, and not in the word candidates selecting experiment.

In the key presses evaluation, the system counts all key presses. We compare a system using the sEMG signals to a system using mouse clicking to indicate what finger is flexed. When using the sEMG based system, the user uses a mouse button only for deleting words, indicating a false recognition.

When evaluating of the average rank of correct word candidate, 144 e-mail messages were used 36 messages at a time. The results of of the hand motion recognition rate evaluation are shown in Table 2; the results of the number of key presses evaluation are shown in Table 3, and the results of the average rank of the correct word candidate evaluation are shown in Table 4.

## 6 Results

Table 2 shows that the average hand motions recognition rate is nearly 80%, with a recognition rate for "key 1" of over 95%. The reason for this high recognition rate is that the model-data for thumb flexion (assigned to "key 1") is very different from the other finger, as can be observed in Table 1. In fact, the flexor pollicis longus (the thumb flexion muscle) is quite distant from the flexor

**Table 2.** Results of the hand motion recognition rate evaluation

	correct recognition	false recognition	recognition rate [%]
1 key	274	10	97
2 key	226	46	83
3 key	205	66	76
4 key	231	120	69
5 key	234	61	79
total	1170	303	79

**Table 3.** Results of the key presses evaluation

	key 1	key 2	key 3	key 4	key 5	enter key	total
sEMG	338	420	467	423	1442	1801	4891
mouse input	239	245	194	153	1018	1647	3496

**Table 4.** Results of the average rank of the correct word candidate evaluation

e-mail	average rank	#words in learning dictionaries
1-36	10.7	103
37-72	7.8	156
73-108	8.18	210
109-144	7.25	274

digitorum superficialis (the muscle for flexing the other fingers). In contrast, the reason for the low recognition rate of "key 4" is that the digitorum for flexing the fourth finger is close to the other digitorums.

Table 3 shows that the number of key presses using the sEMG signals ( $I_{sEMG}$ ) is about 1.4 times higher than while using a mouse button ( $I_{button}$ ). These numbers include user mistakes. If  $I_{button}$  is seen as the number of key presses required when the recognition rate of the sEMG signals is 100%, the recognition rate of all inputs using the sEMG signals ( $p$ ) can be calculated as follows:

$$p = \frac{I_{button}}{I_{sEMG}} \quad (3)$$

In our experiment  $p$  is about 0.71. Thus the input rate of the system using sEMG is about 70% of the achievable rate. This rate is nearly 10% lower than the recognition rate of hand motions shown in Table 2. The presumed reason for this is that the frequency of using "keys 2-4" when selecting word candidates is larger than the frequency when selecting Japanese syllables, and the recognition rate of "keys 2-4" is lower than that for the other keys.

Table 4 shows that the average rank of the correct word candidate generally decreases, as the number of words in the learning dictionary increases. Therefore we can conclude that the learning dictionary is learning effectively. The system can display up to four word candidates at a time, so a user can usually find the correct input word in the first two candidate lists (on average).



## 7 Conclusions

We proposed an input system for mobile terminals using sEMG signals. It uses a method that recognizes the flexion of five fingers and uses these as six input "keys". Even with only six keys, the user can still select Japanese syllables effectively using the Number-Kanji Translation method. The system uses two dictionaries in order to solve the problem of increased number of word candidates that arises from the Number-Kanji Translation method.

To evaluate the system we measure the hand motion recognition rate, the number of key presses, and the average rank of the correct word candidate. The evaluation experiment shows that the average recognition rate of hand motions is almost 80% while the number of key presses using the sEMG signals is about 1.4 times higher than while using a mouse button. It also shows that the system learns word usage effectively, since the average rank of the correct word candidate generally decreases.

In the near future, we will experiment with large documents, will make improvements in the EMG signal processing module, and will modify the system to be capable to effectively use linguistic information even when the EMG signal process recognition fails.

## References

1. Tanaka-Ishii, K., Inutsuka, Y., Takeichi, M.: Japanese input system with digits –Can Japanese be input only with consonants? In: Human Language Technology Conference, San Diego, USA, pp. 211–218 (2001)
2. Tanaka-Ishii, K., Inutsuka, Y., Takeichi, M.: Entering text using a four button device. In: The 19th International Conference on Computational Linguistics, Taipei, Taiwan, pp. 988–994 (2002)
3. T9. T9 text input home page, <http://www.t9.com>
4. Yoshida, H., Ujiie, H., Ishimura, K., Wada, M.: The estimation of muscle fatigue using chaos analysis. *Journal of the Society of Biomechanisms* 28(4), 201–212 (2004) (in Japanese)
5. Suematsu, S., Yokoi, H.: A motion generating system for multi-fingered myoelectric hand. *International Congress Series* 1291, 257–260 (2006)
6. Calhoun Gloria, L., McMillan Grant, R.: Hands-free input devices for wearable computers. In: *Proceedings of Fourth Annual Symposium on Human Interaction with Complex Systems*, pp. 118–123. IEEE, Dayton (1998)
7. Costanza, E., Inverso, S.A., Allen, R.: Toward subtle intimate interfaces for mobile devices using an EMG controller. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, Portland, Oregon, USA (2005)
8. Aso, S., Sasaki, A., Hashimoto, H., Ishii, C.: Driving Electric Car by Using EMG Interface. In: *IEEE International Conferences on Cybernetics Intelligent Systems (CIS)*. IEEE, CD-ROM (2006)
9. Hatano, A., Araki, K., Masafumi, M.: A Japanese Input Method Using Surface EMG Signals. In: *Hokkaido-section Joint Convention of the Institutes of Electrical and related Engineers*, Sapporo, Japan, pp. 187–188 (2007)
10. Kudo, T., Kazawa, H.: Web Japanese N-gram Version 1. Gengo Shigen Kyokai

11. Oisaka Electronic Device Ltd. equipment: Personal-EMG,  
<http://www.oisaka.co.jp/P-EMG.html>
12. Matsuhara, M., Araki, K., Tochinai, K.: Evaluation of Number-Kanji Translation Method using Inductive Learning on E-mail. In: Proceedings of the IASTED International Conference on Artificial Intelligence And Soft Computing, Banff, Canada, pp. 487–493 (2000)
13. Matsuhara, M., Araki, K., Tochinai, K.: Effectiveness for machine translation method using inductive learning on number representation. In: McKay, B., Slaney, J.K. (eds.) Canadian AI 2002. LNCS, vol. 2557, pp. 648–659. Springer, Heidelberg (2002)