Smart Human Computer Interface with EMG and Vision Based on Multi-modal Information Fusion

Hee-su KANG, Hyun-chool SHIN

(*Dept. of Electronic Engineering, School of IT, Soongsil University, Seoul* 156-743, *Korea*)

Abstract – **A smart Human-Computer Interface (HCI) replacing conventional mouse interface is proposed. The interface is able to control cursor and command action with only hand. Four finger motions (left click, right click, hold, drag) are used to command the interface. Also the authors materialize cursor movement control using image processing. The measure what they use for inference is entropy of Electromyogram (EMG) signal, Gaussian modeling and maximum likelihood estimation. In image processing for cursor control, they use color recognition to get the center point of finger tip from marker, and map the point onto cursor. Accuracy of finger movement inference is over 95% and cursor control works naturally without delay. They materialize whole system to check its performance and utility.**

Key words* – *EMG*: *vision*: *HCI*: *interface*: *mouse

Manuscript Number: 1674-8042(2011)02-0152-05

dio: 10.3969/j.issn.1674-8042.2011.02.13

1 Introduction

In these days, computer has been developed and its interfaces are also being developed. Most commonly used interfaces are keyboard and mouse. Conventional mouse is moved perpendicular to the display such as a desk so; there are constraints on use environments. To overcome such shortcoming, various studies are conducted. Recently, touch screen is widely used because; it is easy to use for non-professional users. User has to touch the screen that is why it is hard to be used on big screen. Additionally there are interface under development based on bio signal which is the primary response of the human will. Serby[1] study p300 which is evoked potential of brain wave. Kim[2] study interfaces based on eye gazed tracking. Ko[3] study about interfaces based on hand gesture recognition through image processing.

Interface which based on P300 works slowly[4]. Interface based on eye gazed, that is uncomfortable to select and cancel command[5] and also there is problem of fatigue.

In this study, to solve above problem and for more intuitive computer control, we developed new computer interface through fusion of EMG signal and Image processing. The interface controls the cursor through image processing. Interface is recognizing specific color of marker and then it is mapping marker to cursor through image processing. And also it is controlling command through EMG signal processing, which is inferring the motion.

The composition of this paper is as follows: In chapter 2, explains how to process the EMG signal to control command. In chapter 3, explains how to process the image from video camera to control the cursor. In chapter 4, shows results of the algorithm in chapter 2 and chapter 3. Finally, in chapter 4, conclude the study.

2 EMG signal processing for interface command

2.1 EMG signal acquisition

For EMG signal acquisition, EMG acquisition equipment (PolyG-A, Laxtha Inc.) and bi-polar Ag-AgCl electrode is used. We use 4 channels of the sensor and the sampling frequency is 256 Hz. We use notch filter to remove 60 Hz and 120 Hz power line noise. We implemented a band type sensor which is easy to wear on wrist, as shown in Fig.1(a).

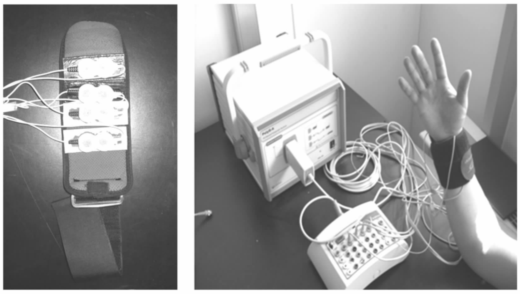
(a) (b)

Fig.1 (a) Wearable 4 channels EMG sensor; (b) Aqcuisition system and wearable EMG Sensor

---

* **Received**: 2010-08-21
Project supported: This work was supported by the MKE (The Ministry of Knowledge Economy), Korea, the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency)(NIPA-2010-C1090-1021-0010)
Corresponding author: Hee-su KANG(uncmincm@gmail.com)

2.2 Model of finger motion for command of interface

We use four finger motions (left click, right click, hold, release) to perform conventional mouse commands. The motions of left click or right click are actual behavior, such as mouse left click or right click, bending index finger or middle finger. Drag and drop which is used for file move in conventional mouse is divided into Drag (move with press down the button) and Drop (stop pressing). As shown in Fig. 2(c) and Fig. 2(d), motions for the file movement are like actually picking something up and releasing it down. Such a configuration of motions helps to command execution with intuition finger motion.

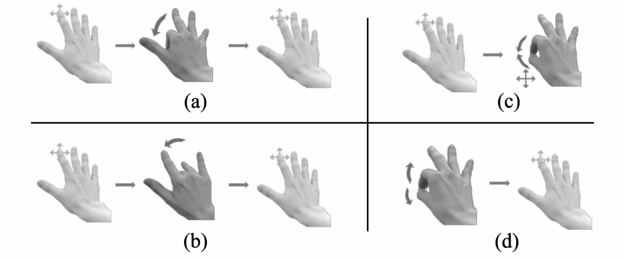

Fig. 2 (a) Finger motion performing left click; (b) Finger motion performing right click; (c) Finger motion performing hold; (d) Finger motion performing release

Fig. 3 shows four channels of EMG signals during each 4 hand motion. Horizontal axis represent time(s), and vertical axis represent voltage(μV) of the signal.

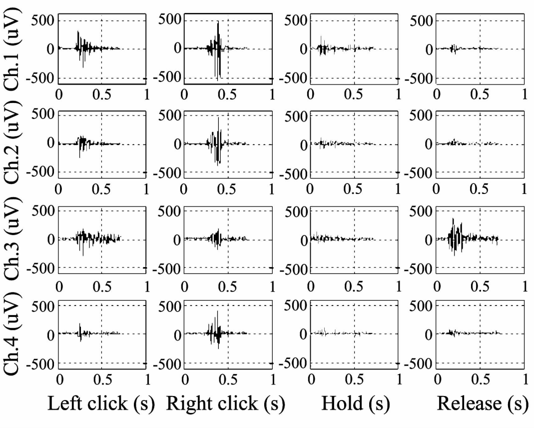

Fig. 3 EMG signals during various hand motions

Left click denoted by the "L", Right click denoted by the "R", Hold denoted by the "H", and Release denoted by the "Z". If any motion $K \in \{L, R, H, Z\}$ happens, signal observed in $N \in \{1,2,3,4\}$ th channel is denoted by the $r_N^K[n]$. For example, If Hold motion happened, the signal from 3rd channel is denoted by $r_4^H[n]$.

2.3 Probabilistic modeling of entropy of EMG signal

We use information entropy to extract feature of the signal. Information entropy is average of self information of random variable. It can measure uncertainty of random variables and imbalance of probability distribution. The entropy is defined as[6]

$$H(W) = -\sum_{i=1}^{L} p(w_i)\log\{p(w_i)\}, \tag{1}$$

where W is $W = \{w \mid w_1, w_2, \cdots, w_L\}$ as discrete random variable. $p(w_1)$ is probability of w_1 and it is satisfied[7]

$$0 \leqslant p(w_i) \leqslant 1,$$
$$\sum_{I=1}^{T} p(w_i) = 1. \tag{2}$$

EMG signals have both negative and positive value, but phase information of the signal is not important, so we process absolute value of the signal. Absolute signal is denoted by $x_N^H[n]$. And

$$x_N^H[n] = | r_N^H[n] |^{\sim}. \tag{3}$$

Let entropy of $x_N^H[n]$ be $H[x_N^H[n]]$. To calculate $H[x_N^H[n]]$, as shown in Eq. (1), probabilities $p(w_i)$ are needed. The probabilities are calculated in the following way.

Divide amplitude segment 0 to $x_{\max}$ into M sections. Individual segments are denoted by $I_1, I_2, \cdots, I_M$. After that, count samples in each segment. For probability modeling, divide number of samples in each segment into the total number of samples.

If motion K happens in Nth channel, probability in I_m is denoted by $p_N^K(m)$.

$$p_N^K(m) = \frac{\#\text{of samples} \in I_m}{\#\ \text{of total samples}},$$
$$I_m = \left\{ n \mid \frac{(m-1)\cdot x_{\max}}{M} \leqslant x[n] \leqslant \frac{m \cdot x_{\max}}{M} \right\},$$
$$m = 1, 2, \cdots, M. \tag{4}$$

where # of samples $\in I_m$ means that number of samples which are contained in I_m. Thus, entropy of signal $x_N^H[n]$ can be expressed as

$$h_N^H = H(x_N^H[n]) = -\sum_{i=1}^{L} p_N^K(m)\log\{p_N^K(m)\}^{\sim}. \tag{5}$$

Next is the processing of creating probability model of entropy which for discrimination. If a command motion $K \in \{L, R, H, Z\}$ happens, entropy, one for each channel occurs. To examine the statistical characteristics of entropy, we performed 100 times for each command motion (L, R, Z, H).

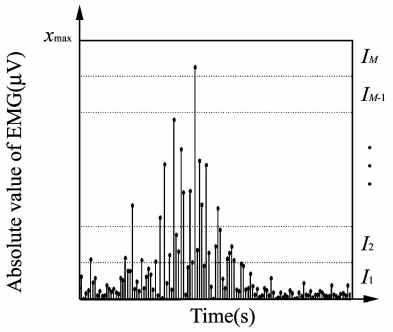

Fig. 4 EMG amplitude division for probabilistic modeling

Fig. 5 shows histogram of entropies for channel and the motion. Horizontal axis represents value of entropy, and vertical axis represents frequency (counts/bin) of entropy value.

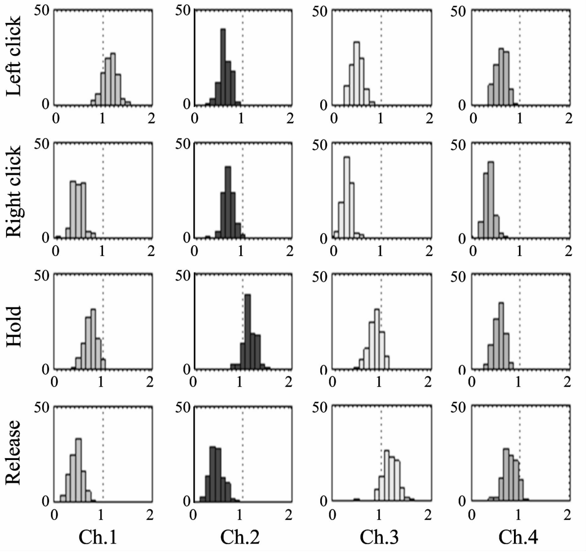

Fig. 5 Histogram of each channel's entropy

Fig. 6 shows probabilistic models which are derived from the histograms in Fig. 5. The reason why do we apply the Gaussian distribution model is figure of histogram from 100 times of trial similar to Gaussian shape. Entropy histogram from training data, through Eq. (6), that can be converted to a Gaussian distribution function $f_N^K(h)$ [7].

$$f_N^K(h) = \frac{1}{\sqrt{2\pi(\sigma_N^K)^2}}\exp\left(-\frac{(h-\mu_N^K)^2}{2(\sigma_N^K)^2}\right), \qquad (6)$$

where $f_N^K(h_N)$ is probability density function from channel N with motion K . h_N is entropy of the signal from channel N , and it is index of the probability density function $f_N^K(h_N)$. Fig. 6 shows that probability of entropies per channel, and the motion. Horizontal axis represents value of entropy, and vertical axis represents probability density of entropy. Through this modeling, we can get the probability density of each motion in arbitrary channel.

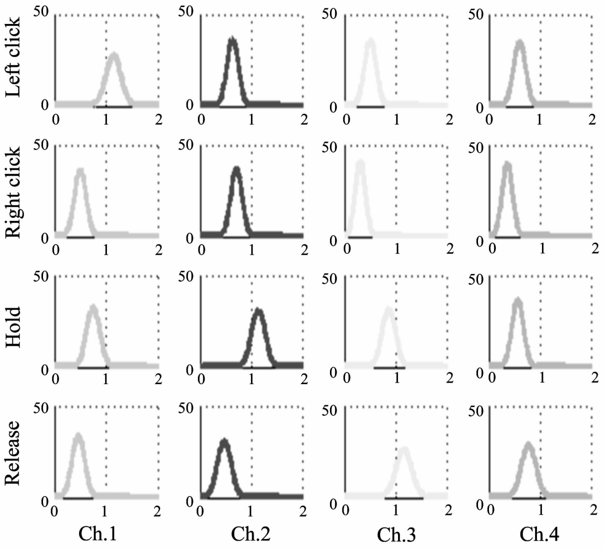

Fig. 6 Gaussian model from the histogram

2.4 Command motion inference with maximum likelihood estimation

Maximum likelihood estimation with probability model of entropy is used for command motion inference. If any motion $K \in \{L,R,H,Z\}$ happened, we can get the probability density function value $(f_N^L(h), f_N^R(h), f_N^H(h), f_N^Z(h))$ through entropy h from N^{th} channel. The values mean probability quantity of entropy h happened, so the value can be interpreted as likelihood[8]. Through maximum likelihood estimation, we infer the command motion $\hat{K}$ which has largest likelihood. Under assumption that functions are statistical independent, Joint probability density function of motions for each channels are can be express as a multiplication of one-dimensional probability density functions. Then, the inferred command motion ($\hat{K}$) is can be expressed by

$$\hat{K} = \arg\max_K \left\{ \prod_{\eta=1}^{4} f_\eta^K(h_\eta^w) \right\}. \qquad (7)$$

3 Image processing for cursor control

3.1 Image acquisition

Hand movement recognition through image processing is used for cursor control of the interface. Input images are acquired from the video camera (Webcam VX-3000, Microsoft). The camera support 640 × 480 resolution images and it can acquire 30 frames per second. It is installed against the user's hand to get the images of hand movement.

3.2 HSV and marker color recognition

When hand movement recognized with image from video camera, a color marker is used for more accurate recognition. The color of the marker on finger tip is not included in the background color. Image data from the video camera are composed of RGB color model. The RGB color model is hard to be used for color recognition, because it is so sensitive to light. To minimize the effects of light, RGB color model is converted to HSV color model which is contained by hue, saturation, and brightness value[9]. The transform is applied by Eq. (8)[10]. Pixels in the region of marker are detected by threshold. Threshold of HSV values were set throughout many experiment as follows:

$$(120 < H < 170)\&(70 < S)\&(140 < V).$$

Through C pixels of the marker which satisfy the above conditions, the center position of the marker (I, J) can be get with equation (9)[10-11].

In Fig. 7, (a) is hand with marker on ring finger, (b) is picture that region of marker is detected so displayed in red, and (c) is picture that shows center of the region.

$$\Phi = \max(R,G,B),$$
$$\varphi = \min(R,G,B),$$

$$H = \begin{cases} 0^\circ, & \Phi = \varphi; \\ 60^\circ \times \dfrac{G - B}{\Phi - \varphi} + 0^\circ, & \Phi = R; \\ 60^\circ \times \dfrac{G - B}{\Phi - \varphi} + 130^\circ, & \Phi = R; \\ 60^\circ \times \dfrac{G - B}{\Phi - \varphi} + 240^\circ, & \Phi = R. \end{cases} \tag{8}$$

$$S = \begin{cases} 0, & \Phi = 0; \\ \dfrac{\Phi - \varphi,}{\Phi} & \Phi \neq 0. ' \end{cases}$$

$$V = \Phi.$$

$$(I, J) = \left(\frac{1}{C} \sum_{k=1}^{C} i_k, \frac{1}{C} \sum_{k=1}^{C} j_k \right). \tag{9}$$

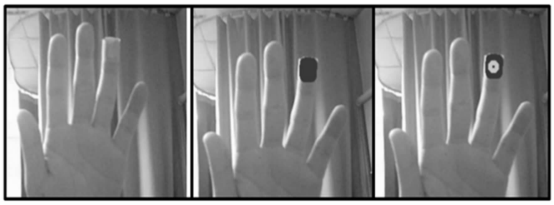
(a) (b) (c)

Fig. 7 (a) Hand with marker in the ring finger; (b) Recognized region of marker; (c) Center point in the region

Hand movement recognition through image processing is used for cursor control of the interface. Input images are acquired from the video camera (Webcam VX-3000, Microsoft). The camera support 640 × 480 resolution images and it can acquire 30 frames per second. It is installed against the user′s hand to get the images of hand movement.

4 Experimental result

For verify the accuracy of inference of each motion, each motion was independently performed 50 times. 50 data are separated by 25 training data and 25 test data. There are ${}_{50}C_{25}$ ($\approx 1.26 \times 10^{14}$) total combinations. To verify the accuracy of motion inferring, we randomly choose 1 000 combinations in total combinations.

To calculate entropy, $X_{\max}$ was set at 1 000 μV. And number of segment M was set at 8. 1000 times of the test, infference accuaracy are shown in Tab. 1.

Tab. 1 Accuracy of motion inference($M = 8$)

Motion	Inference			
	L(%)	R(%)	H(%)	Z(%)
Left click	99.37	0.00	0.63	0.00
Right click	0.00	96.04	3.96	0.00
Hold	4.42	0.00	92.54	3.04
Release	0.00	0.00	0.13	99.87

5 Conclusion

Fig. 8 shows that average accuracy depending on parameter M. and Fig. 9, Fig. 10 and Fig. 11 show that how does the interface works. Fig. 9 shows right click command with middle finger bending. Fig. 10 shows left click command with index finger bending. Fig. 11 shows drag and drop command through finger motion which hold actual something.

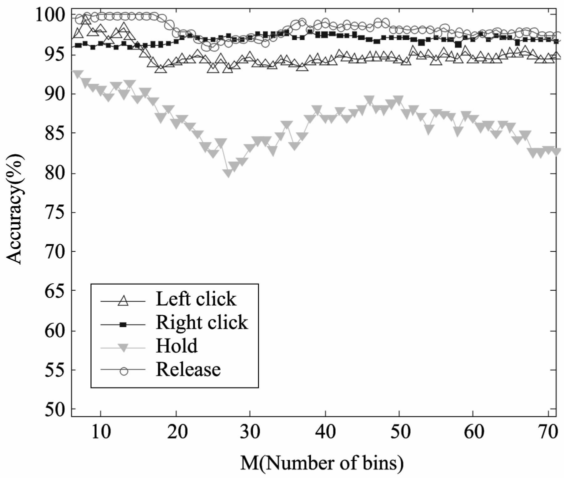

Fig. 8 Accuracy depending on parameter M

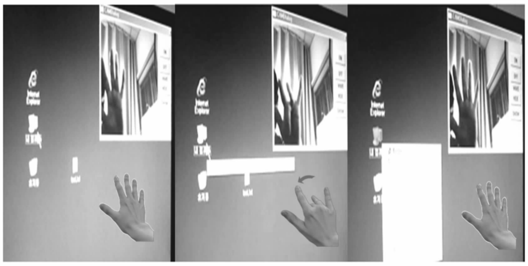

Fig. 9 Middle finger motion for right click command

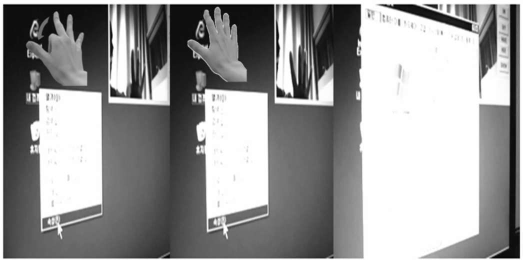

Fig. 10 Index finger motion for right click command

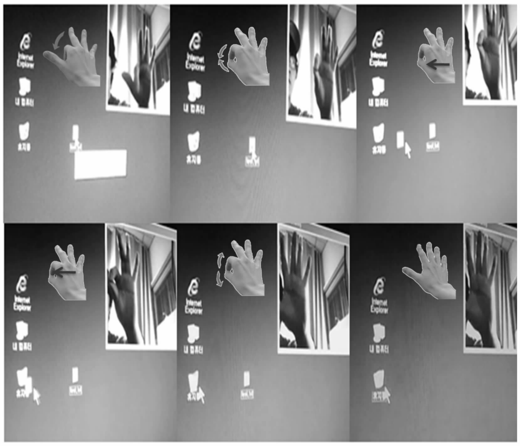

Fig. 11 Drag and frop command with holding gesture

We developed new human computer interface through fusion of image processing and EMG signal processing. As we deal with earlier, the interface control not only cursor movement but also command alike conventional mouse.

In hand movement recognition, marker color was recognized to find region of marker and center of the region was mapped to cursor. Only single color was used for the marker. So if background color is similar to marker, center of the region was located outside of marker, so the cursor movement was swayed. We expect that problem will be solved through using two or more colors of marker.

Accuracy of finger motion inferring is, as shown in Tab.1, more than 95%. Relation of motion “H” and “Z” for drag and drop, Z (release) could not happen before H (hold). So, if we prevent release motion inferring before hold happen, we can correct wrong identification.

In this study, we developed computer interface which is capable of 4 commanding through EMG sensor and video camera. If we use more EMG sensor and improved image processing algorithm, then this interface system will be developed with various functions.

References

[1] H. Serby, E. Yom-Tov, G. F. Inbar, 2005. An improved P300-based brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 13(1): 89-98.

[2] K. N. Kim, R. Ramakrishna, 1999. Vision-Based Eye-Gaze Tracking for Human Computer Interface. International Conference on Systems, Man, and Cybernetics, p. 324-329.

[3] B. K. Ko, H. S. Yang, 1997. Finger mouse and gesture recognition system as a new Human computer interface. *Computers & Graphics*, 21(5): 555-561.

[4] E. Donchin, K. Spencer, R. Wijesinghe, 2000. The mental prosthesis: assessing the speed of a P300-basedbrain-computer interface. *IEEE Transactions on Rehabilitation Engineering*, 8 (2): 174-179.

[5] T. Yagi, 2006. Drifting and Blinking Compensation in Electro-Oculography (EOG) Eye-Gaze Interface. IEEE International Conference on Systems, Man and Cybernetics, SMC′06.

[6] R. Bose, 2003. Information Theory, Coding and Cryptography. Tata McGraw-Hill.

[7] A. Papoulis, 2002. Probability, Random Variables, and Stochastic Processes. McGraw-Hill, New York.

[8] KJ. You, HC. Shin, 2009. Classifying finger flexing motions with surface EMG using entropy and the Maximum likelihood method. *Jounal of IEEK-SC*, 46: 38-43.

[9] K. H. Moon, 2006. Hybrid Color Model for Robust Detection of Skin Color. KIISE Conference, p. 98-101.

[10] M. S. Ko, 2009. An implementation of user interface using vision-based gesture recognition. *Journal of KIISE: Computer Systems and Theory*, 35(1): 507-511.

[11] J. W. Son, 2009. Extraction of color information from images using grid kernel. *Journal of KIISE: Computer Systems and Theory*, 34(1): 182-187.