
4 Social Interaction

4-1 Basic Study for Cognition and Manipulation of the Body Image

MAEKAWA Satoshi

In recent years, it becomes clear that the body image is not inherent, but have plasticity. This fact suggests that the body image can be manipulated. With the progress of the computer, virtual external world can be built with reality, but the ability of the body image manipulation suggests virtual self with reality. Although the cognitive ability of human is limited, the world of virtual self with reality, which means freedom from physical body, may be very vast. However, the comprehension level about virtual self is just low so far, and a basic study of it is conducted as before. In this paper, estimation methods of amputee's motor intention from surface electromyography measured noninvasively, are proposed, and the results of psychological experiments about body image are shown.

Keywords

Body image, Surface EMG, Motor Unit, Mirror neuron

1 Introduction

In recent years, studies have begun to reveal that an individual's conception of his or her body image is not inherent and instead features notable plasticity[1]. It is therefore possible in principle to manipulate body image. Progress in computer technology has allowed us to construct an ultra-realistic virtual world that closely reflects the real world; however, the manipulation of body image must go a step beyond this virtual world to enable the creation of a realistic virtual self. Although there are limits to human cognitive ability, such virtual body imaging may open the door to a vast and unprecedented world free from the restrictions of the physical body. However, at present, knowledge of the virtual self is still limited; we are still at the stage of basic research. In the present paper, we will propose a non-invasive detection method for motor

intention by surface electromyography that will enable us to investigate the phantom limb phenomenon found in amputees, and will introduce the results of a basic psychological experiment conducted to study the cognitive processes involved in a body image with plasticity. A study of the decomposition of motor units — the minimum units of muscle movement from multi-channel surface electromyography — will be introduced in Section 2, and our basic understanding of body image cognition to date will be introduced in Section 3.

2 Decomposition of motor units based on surface electromyography

2.1 Surface electromyography

The motor unit (MU) is the minimum functional unit of muscles, and consists of a single α motor neuron within the spinal cord

and a group of muscle fibers controlled by the neuron. The contraction of a muscle generates MU action potentials (MUAPs), which can be measured to study the activity of MUs. However, muscle activity normally involves simultaneous activation of multiple MUs. Therefore, the observed electromyography (EMG) will be the result of the superposition of all MU activity. Thus, the decomposition of MU activity is essential to a complete understanding of muscle activity, and — along with the identification of MU activity — forms an important element of current research.

Surface EMG (sEMG) measures the temporal changes in electric potential produced on the surface of the skin due to the migration of local depolarization in the muscle fiber within body tissue at speeds of 2-6 m/s. Generally, a point of depolarization initially positioned at a distance gradually approaches the electrode, passes in proximity to the instrument, and moves away. This process corresponds to measurement of sEMG through a non-causal filter. When observed as an impulse response, this process is seen as a gradual increase in amplitude to the peak of the signal, followed by a gradual attenuation.

To check the general characteristics of the sEMG signal, we will first examine a synthe-

sized virtual sEMG signal. For simplicity, a semi-infinite conducting region with planar boundaries is assumed. Within this conductive body, a current dipole 1 mm in length is moved at 3 m/s at a depth of 1 cm. The electric potential induced by the dipole will be measured as it passes directly below two electrodes placed 1 cm apart on the surface of the conductive body. The sampling frequency is set at 1 kHz. Figure 1 (left) shows the change in electric potential observed in the signal for an adequate time frame. The actual duration of the impulse response depends on the position of the motor endplate and the tendon, but here, this period was assumed to be approximately 50 ms, equivalent to an assumed muscle-fiber length of approximately 150 mm.

Figure 1 (right) shows the pole and zeros determined using this signal as the impulse response. It can be seen from the figure that the zeros fall outside of the unit circle. In other words, the sEMG signals were found to be non-minimum phases.

Since the non-minimum phase filter has zeros outside of the unit circle, we must make allowances for a phase lag when making calculations for the inverse filter, or deconvolution. Therefore, in order to extract MU from sEMG, a decomposition method must be

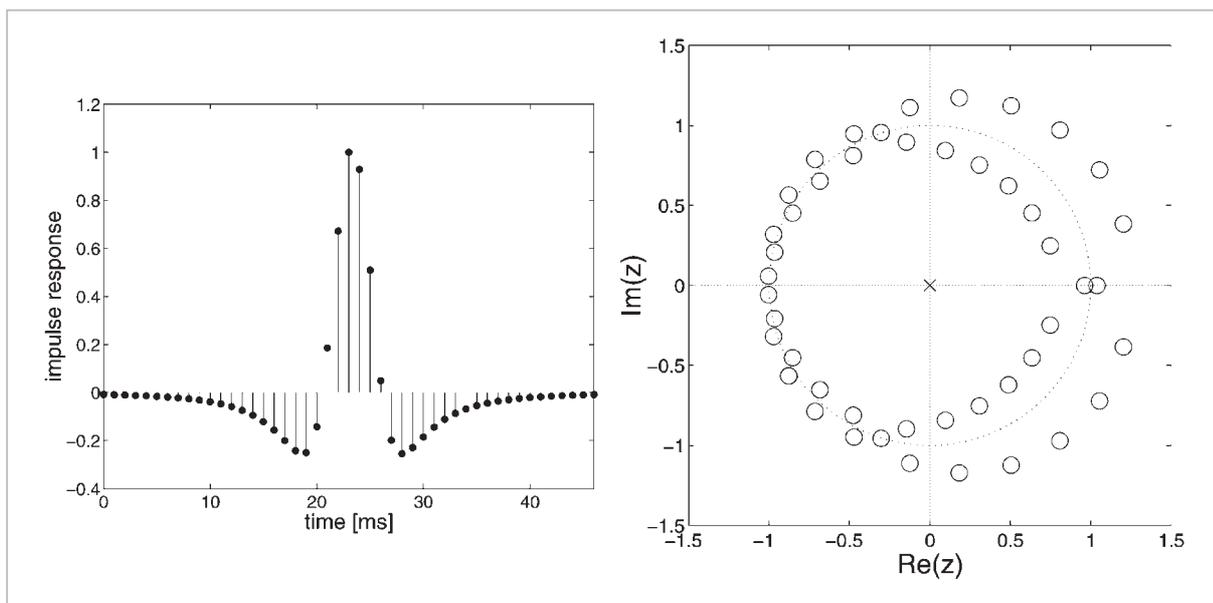


Fig. 1 Impulse response (left) and location of the pole (cross) and zeros (open circles) (right) of virtual surface EMG

applied that can handle non-minimum phase characteristics.

2.2 MU decomposition by blind deconvolution^[3]

A method of blind deconvolution for non-minimum phase characteristics has been proposed by Zhang et al.[2], and we will attempt to approach MU decomposition using this method. The procedure is as outlined below. When it is assumed that an n number of time-series signals, $\mathbf{s}(t) = [s_1(t), \dots, s_n(t)]^T$, statistically independent in time and space, are passed through a mixed filter $\mathbf{A}(z^{-1})$ having non-minimum phase characteristics, and that the resulting signal observed for time t is represented by $\mathbf{x}(t) = [x_1(t), \dots, x_n(t)]^T$, then the multi-channel blind deconvolution method will decompose the signal into independent components $\mathbf{y}(t) = [y_1(t), \dots, y_n(t)]^T$ using the following relationship.

$$\mathbf{y}(t) = \mathbf{W}(z^{-1})\mathbf{x}(t) \quad (1)$$

Here, z^{-1} is the time-lag operator. To address the non-minimum phase characteristics, the non-causal filter $\mathbf{W}(z^{-1})$ is represented by

$$W_{ij}(z^{-1}) = \sum_{\tau'=-\tau}^{\tau} W_{ij}(\tau')z^{-\tau'}. \quad (2)$$

Based on the above, the independent component $\mathbf{y}(t)$ may be expressed as follows.

$$\mathbf{y}(t) = \mathbf{W}(z^{-1})\mathbf{A}(z^{-1})\mathbf{s}(t) \quad (3)$$

When $\mathbf{W}(z^{-1})\mathbf{A}(z^{-1})$ is equivalent to the unit matrix, then $\mathbf{y}(t) = \mathbf{s}(t)$, and the signal may be fully reproduced. However, in reality, uncertainties remain due to the symmetrical properties of the order of the indexes, scales of each component, and the time lag.

Next, in order to simplify the calculations for determining the inverse filter, $\mathbf{W}(z^{-1})$ is broken down into two single-sided FIR filters \mathbf{L} (causal) and \mathbf{R} (non-causal). The mixed filter is assumed to be temporally invariable, and the learning of the inverse filter is performed according to

$$\Delta\mathbf{L}(\tau) = \eta \sum_{\tau'=0}^{\tau} \left\{ \delta(\tau - \tau') - \varphi(\mathbf{y}(t))\mathbf{y}^T(t - \tau') \right\} \mathbf{L}(\tau - \tau') \quad (4)$$

$$\Delta\mathbf{R}(\tau) = -\eta \sum_{\tau'=0}^{\tau} \mathbf{L}^H(z)\varphi(\mathbf{y}(t))\mathbf{u}^T(t + \tau)\mathbf{R}(\tau - \tau') \quad (5)$$

$$(\tau = 0, \dots, N).$$

Note that $\mathbf{R}(0)$ is a unit matrix, and so learning does not take place. Here, η is the learning coefficient, $\varphi(\mathbf{y})$ is the non-linear function, and $\mathbf{L}^H(z) = \sum_{\tau=0}^N \mathbf{L}^T(\tau)z^{\tau}$.

2.3 MU decomposition by overcomplete representations

The blind deconvolution method introduced in the previous section is a separation method based on linear calculations, and thus cannot handle cases in which the number of signal sources exceeds that of the observed channels. MU decomposition with overcomplete representation is used to overcome this problem. To apply this process to the decomposition of multi-channel time-series signals, it is assumed that identical bases — shifted by one-sample intervals in the time axis direction — are placed throughout the target time frame[7].

The derivation process for overcomplete representation is reported in detail in Lewicki et al.[4]-[6], and so only a brief summary will be given here. It is assumed that the L -dimensional observed signal \mathbf{x} may be expressed as follows using the M -dimensional signal source \mathbf{s} .

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \boldsymbol{\varepsilon} \quad (6)$$

Here, \mathbf{A} is a $L \times M$ base matrix. In order for this matrix to be overcomplete, $L < M$, so even if \mathbf{A} were defined, the signal source \mathbf{s} cannot be determined straightforwardly for the observed signal \mathbf{x} . Note that $\boldsymbol{\varepsilon}$ is Gaussian noise and $\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T = \lambda^{-1}\mathbf{I}$. Based on the above, the probability of observation of observed signal \mathbf{x} for a given base and signal can be expressed as

$$P(\mathbf{x} | \mathbf{A}, \mathbf{s}) \propto \exp\left[-\frac{\lambda}{2}\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}\right], \quad (7)$$

since $\boldsymbol{\varepsilon} = \mathbf{x} - \mathbf{A}\mathbf{s}$. The purpose of the overcomplete representation analysis is to estimate the

most plausible base \mathbf{A} and signal source \mathbf{s} from the observed signal \mathbf{x} .

First, we will deal with the estimation of the signal source \mathbf{s} . When $P(\mathbf{s})$ is the a priori probability for \mathbf{s} and $\mathbf{R}(\mathbf{s}) = \frac{\lambda}{2} \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}$ and $\mathbf{S}(\mathbf{s}) = -\log P(\mathbf{s})$, the estimated value $\hat{\mathbf{s}}$ may be given by maximum a posteriori probability (MAP) as

$$\begin{aligned} \hat{\mathbf{s}} &= \arg \max_{\mathbf{s}} \log P(\mathbf{s} | \mathbf{A}, \mathbf{x}) \\ &= \arg \min_{\mathbf{s}} [\mathbf{R}(\mathbf{s}) + \mathbf{S}(\mathbf{s})]. \end{aligned} \quad (8)$$

Since a super Gaussian distribution having statistically independent variables ($P(\mathbf{s}) = \prod_i P(s_i)$) may be given as $P(\mathbf{s})$, the a priori probability of \mathbf{s} , $\mathbf{S}(\mathbf{s}) = -\sum_i \log P(s_i)$, may be regarded as an indicator of sparseness. Moreover, taking the fact that \mathbf{R} is the squared error of reconstruction into consideration, $\hat{\mathbf{s}}$ may be regarded as the estimated signal source from which the observed signals may be reconstructed using the sparsest expressions possible. Note that, in practice, the gradient method may be utilized for determining $\hat{\mathbf{s}}$, and that it is sufficient to solve the equation

$$\frac{d\mathbf{s}}{dt} = \lambda \mathbf{A}^T \boldsymbol{\varepsilon} - \boldsymbol{\varphi}(\mathbf{s}). \quad (9)$$

Here, $\boldsymbol{\varphi}(\mathbf{s}) = -\nabla \log P(\mathbf{s})$.

Next, we will present the learning algorithm for determining the most suitable base for the data structure. Here, the a posterior probability for a given data item \mathbf{x} is maximized based on the maximum likelihood estimation method in order to determine the most likely base \mathbf{A} by the gradient method. The objective function is given by the following equations.

$$L = E[\log P(\mathbf{x} | \mathbf{A})] \quad (10)$$

$$P(\mathbf{x} | \mathbf{A}) = \int P(\mathbf{x} | \mathbf{A}, \mathbf{s}) P(\mathbf{s}) d\mathbf{s} \quad (11)$$

Here, Gaussian integration is applied by approximating the Gaussian distribution for $P(\mathbf{x} | \mathbf{A}, \mathbf{s}) P(\mathbf{s})$ at $\mathbf{s} = \hat{\mathbf{s}}$, and the learning equation is determined by differentiating the log likelihood L by \mathbf{A} . Furthermore, by taking the natural gradient into consideration, the learn-

ing equation below is ultimately obtained.

$$\Delta \mathbf{A} = \eta \mathbf{A} (\boldsymbol{\varphi}(\hat{\mathbf{s}}) \hat{\mathbf{s}}^T - \lambda \mathbf{A}^T \mathbf{A} \mathbf{H}^{-1}(\hat{\mathbf{s}})) \quad (12)$$

Here,

$$\mathbf{H}(\mathbf{s}) = -\lambda \mathbf{A}^T \mathbf{A}^{-1} - \nabla \nabla^T P(\mathbf{s}) \quad (13)$$

where η is the learning coefficient. Note that when λ is sufficiently large, the following approximation may be made.

$$\Delta \mathbf{A} = \eta \mathbf{A} (\boldsymbol{\varphi}(\hat{\mathbf{s}}) \hat{\mathbf{s}}^T - \mathbf{I}) \quad (14)$$

2.4 Application to surface EMG^[8]

(1) Measurement of multi-channel sEMG

The two methods described above were applied in practice to the measured multi-channel sEMG [Fig. 2 (left)] in an attempt to estimate MU. The 20,000-point data set, formed of observations at intervals of 1 ms using a 16-channel sEMG to assess the generation of tensile force in the fourth finger, was used as the observed signal. Fig. 2 (right) presents the observed signals for 1,000 ms excerpted from the entire data set.

(2) MU decomposition by complete representation

Figures 3 (left) and 3(right) present examples of the results of the blind deconvolution method described in Section 2.2 and separation by the overcomplete representation analysis described in Section 2.3, respectively. In both examples, separation was performed using all 16 channels, and three estimated source signals for which clear outputs appeared were selected. Both feature successfully separated signals believed to belong to 3 MUs. The conclusion that these three signals belonged to three different MUs is consistent with statistical observations, such as the distribution of amplitude and firing intervals.

(3) MU decomposition by overcomplete representation

The results of overcomplete representation analysis are shown in Fig. 4. Here, the observed signals were limited to those from 2 channels to create an overcomplete condition. The channels selected were ch 5 and ch 12, which were positioned far apart and

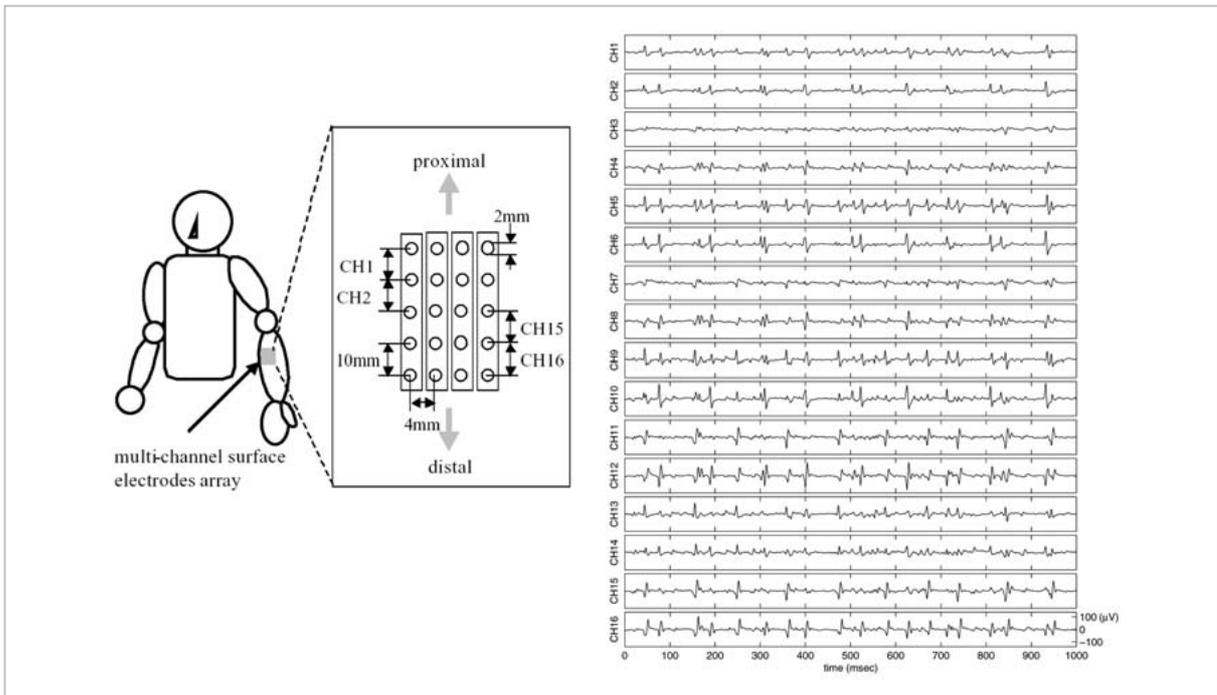


Fig.2 Electrode position (left) and observed signal (right)

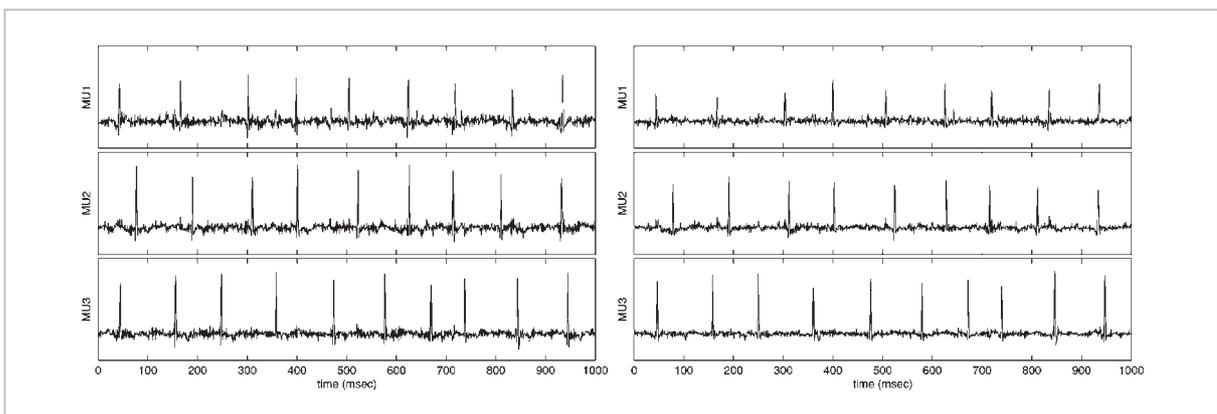


Fig.3 Results of decomposition by blind deconvolution (left) and overcomplete representation analysis (right)

both of which featured large signal power. In this example, the two MUs seem at first glance to be separated relatively clearly. However, upon closer inspection it may be seen that separation is incomplete for MU 1. The coefficient of correlation to the separated signal MU 1 in Fig. 3 (right) is 0.33; a certain degree of positive correlation was therefore confirmed. For artificially synthesized signals, it has been confirmed that three source signals may be successfully estimated even with two-channel observation. This suggests that MU decomposition may be performed more clearly through

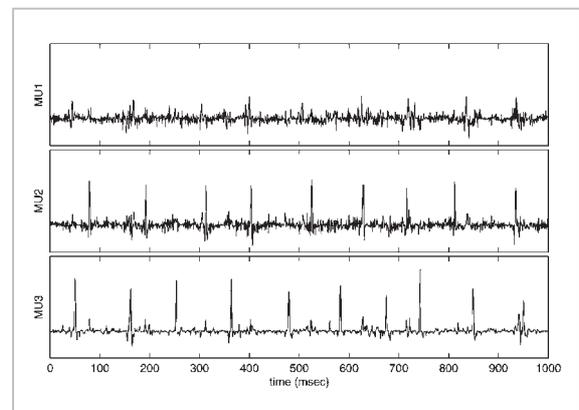


Fig.4 Three source signals separated from the two-channel observation signal

selection of a certain number of observation channels.

3 Cognition of body image^[9]

3.1 Cognition with fingers crossed

An experiment was conducted in which participants were asked to move the finger to which tactile stimulation was applied, under two different sets of conditions: one in which the fingers of both hands were placed over each other, palms-down, so that the fingers on both hands crossed (“crossed state”), and one in which both hands were simply placed side by side (“uncrossed state”), both for cases in which the hands were visible and invisible to the subjects (Fig. 5). The results shown in

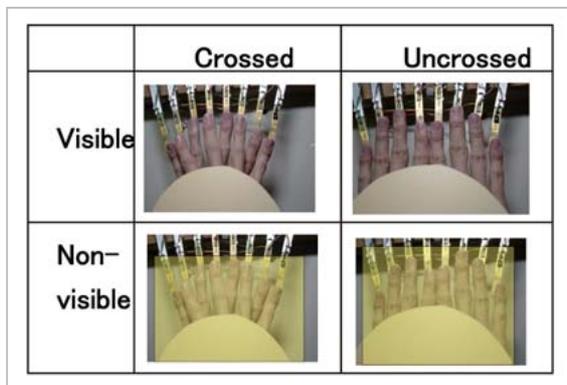


Fig.5 Exercises in tactual stimulation differentiation for crossed and uncrossed fingers

Fig. 6 indicate that although most errors were ipsi-lateral — meaning that the motor response occurred on a different finger in the same hand as the stimulated finger — there were also many cases of contra-lateral hand errors in which the motor response occurred on a finger on the other hand. Furthermore, while the ipsi-lateral errors appeared to be unrelated to posture or visibility, the tendency of contra-lateral errors increased with crossed fingers. Visibility also seemed to enhance, although only slightly, the probability of contra-lateral error.

The task in the present experiment required the tactually stimulated finger to be moved, so this task may be considered to have extremely high SR compatibility, which means that only somatosensory information is essential for its execution. However, in contrast to such expectations, the contra-lateral error occurring with the fingers crossed implies that error may be made in spatial vicinities within space coordinates where the fingers are crossed. The results of this task lead us to the conclusion that not only somatosensory coordinates but also space coordinates are employed in such motor response.

Furthermore, although small, the effect of body visibility is notable. With subjects’ fingers crossed, contra-lateral error increases even without body visibility, and we therefore

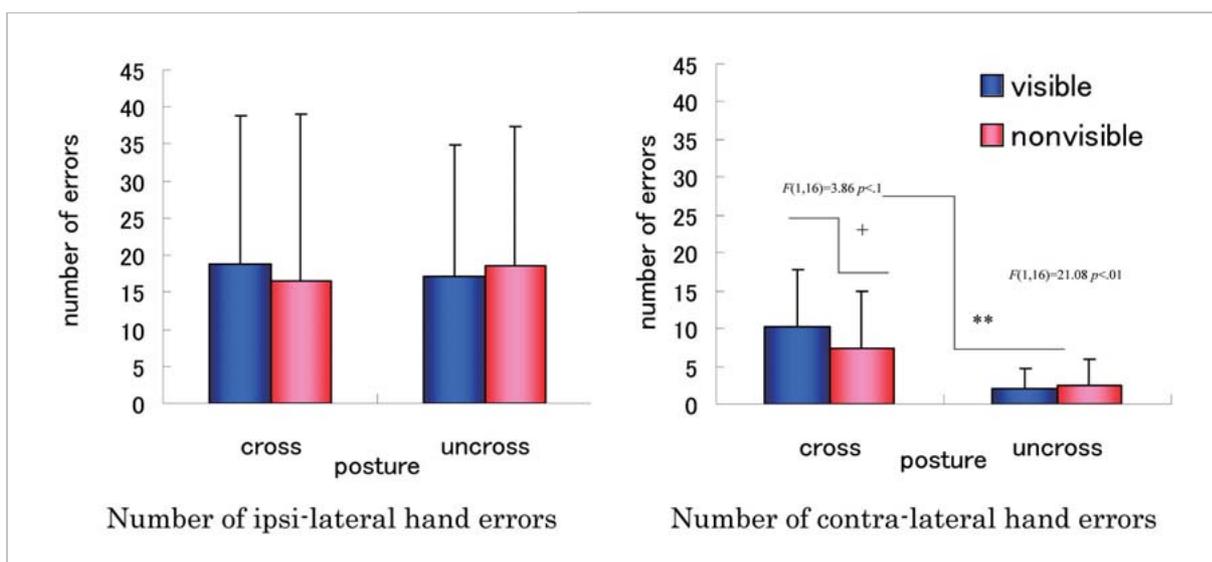


Fig.6 Number of errors in tactual finger discrimination task with crossed and uncrossed fingers

conclude that proprioception is sufficient for the construction of body image in space coordinates.

3.2 The effects of body visibility

In the present section we examine the effects of body visibility on body image cognition. Figure 7 presents the various visual sensation conditions employed: (1) actual body view, (2) a photograph resembling the actual posture, (3) a line drawing resembling the actual posture, and (4) a neutral object. The sense of reality is expected to decrease from the actual body view through the remaining items, in order. The exercise employed the same tactile discrimination scheme described in Section 3.1.

Figure 8 shows the experimental results for contra-lateral errors. These results show a maximum degree of difference between the actual body view (1) and the photographed view (2), which views presumably present the

minimum difference in visual stimulation. There were no statistical differences between the results with an actual body view (1) and a neutral object (4), which corresponds to the maximum difference in visual stimulation. These results seem puzzling, with differences observed for similar situations but not for vastly different conditions.

In the absence of any model to explain such phenomenon, we will propose a hypothesis beginning from certain reports — involving so-called “mirror neurons” — that may prove relevant [10]-[12]. However, the mirror neurons in these studies are explained as neurons that perform coding of an identical motion between the self and a non-self other, leading to the conclusion that such neuronal mechanisms should, strictly speaking, be distinguished from the mechanisms of static posture phenomena (as in the present experiment). Nevertheless, the presence of some mirror-neuron-like neuron that fires for an identical static posture — regardless

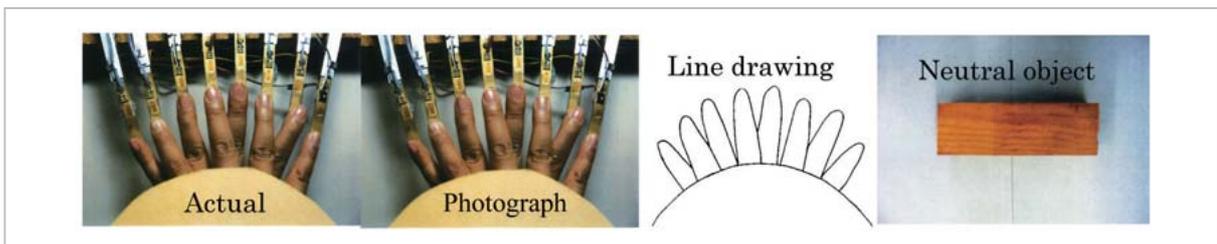


Fig.7 Four conditions for examining the effects of body visibility

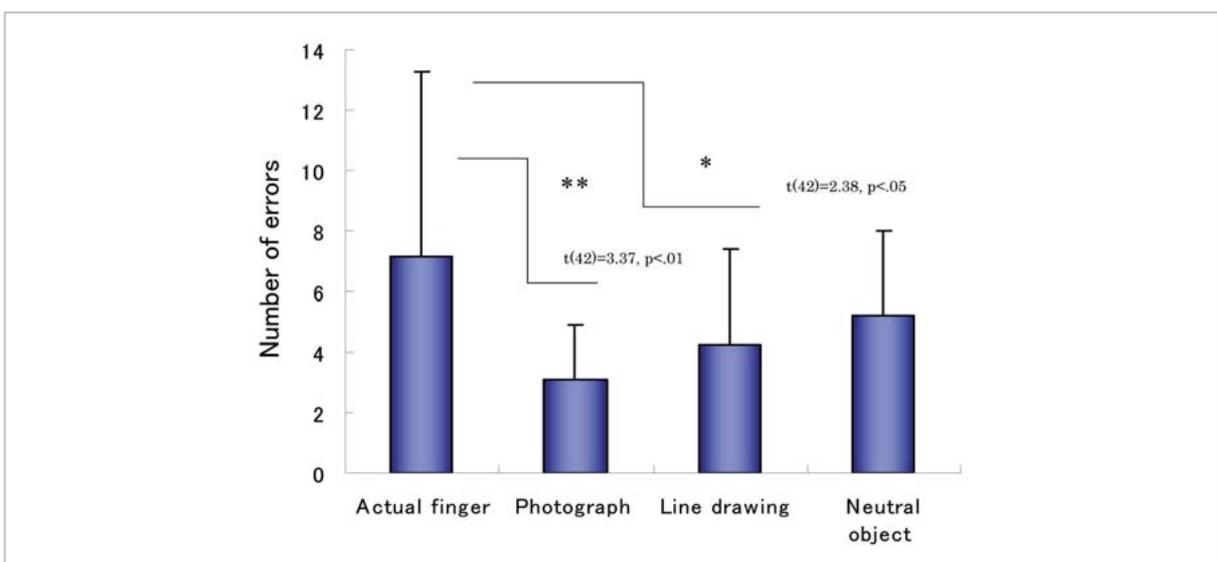


Fig.8 Number of contra-lateral errors for the respective visibility states

of whether the identical posture is of the self or the non-self — may explain the above phenomenon quite well, as described further below.

The difference between (1) and (2) lies in the question of whether or not minute motion in the fingers can be perceived. Thus, the participant is able to see the minute motion associated with the motor command of the self in state (1), while such motion cannot be observed at all in state (2), even though the motor command has been sent. This may result in registration of the body image visually presented in state (2) as a non-self image, even though it does not visually conflict with the actual body self.

Let us assume here that a certain neuron is responsible for the coding of space coordinates for the crossed-finger posture, albeit with differences between self and non-self situations. As described in Section 3.1, this neuron should be responsible for tactile finger discrimination under both directly visible and completely blind conditions. In such a case, what would happen if this neuron were to be used for the non-self, like the photographed view in state (2)? Simultaneous neuron firing will be involved in determining the self/non-self attribution. In other words, binding will occur through simultaneous firing. In the non-self firing, binding may not take place between the processing paths executing the finger discrimination task. This means that there will be no intervention in the processing system of space coordinate coding by the neuron, and as a result the number of contra-lateral errors will be reduced.

The above may be summarized as follows. First, a certain neuron performs coding of space coordinates for a static posture. This neuron will fire for both proprioception and visual body image. When a photograph containing identical posture information is provided as the visual body image, this neuron will be active with visual dominance, and will code the viewed posture as non-self. In other words, the resources regarding the self/non-self (specifically, the neurons that code a given static posture in space coordinates) will compete

with one another, and the coding resource will be used as the non-self posture, since this code is given priority by visual dominance. As a result, the processing system for finger discrimination will only be able to use somatosensory information, and so contra-lateral error is reduced.

The results of the present experiment may also be interpreted as follows. Although there is initially almost no difference between states (1) and (4), the addition of the photograph in state (2) results in the suppression of contra-lateral errors, whose occurrence normally would not have conflicted with the given conditions due to its proximity in the visual map. At first glance, this phenomenon seems in direct opposition to visual dominance, a characteristic well known in psychology. However, based on the interpretation provided above, the phenomenon may be understood as a suppression mechanism caused by competition among resources resulting from visual dominance, and thus does not conflict with past findings.

4 Conclusions

In the present paper it was first shown that surface EMG has non-minimum phase characteristics, and that motor unit decomposition may be performed by applying a multi-channel blind deconvolution method that can handle such characteristics. Further, overcomplete representation analysis may be expanded for application to multi-channel time-series signals by providing bases with symmetry for shifting along the time axis, thus enabling application in sEMG. Results indicate that it may be possible to extract MU featuring a number of channels in excess of the number of those observed.

We have also conducted a psychological experiment on the effects of body visibility on body-image construction. The effects of self/non-self attribution on the processing path were discussed and a new model proposed.

The present research is a part of the Grant-in-Aid for Scientific Research (B)(2), “Research on the Visualization of Motor Unit

References

- 1 V. S. Ramachandran and S. Blakeslee, “Phantoms in the brain: Probing the mysteries of the human mind”, New York, William Morrow & Co, 1998.
- 2 L. Zhang, A. Cichocki, and S. Amari, “Multichannel blind deconvolution of non-minimum phase systems using information backpropagation”, Proc. of the 5th Int. Conf. on Neural Information Processing, pp.210-216, 1999.
- 3 S. Maekawa, T. Arimoto, and M. Kotani, “MU decomposition from multichannel surface EMG signals using blind deconvolution”, Electronics and Communications in Japan, Part 3, Vol.90, No.2, pp.22-30, 2007.
- 4 M. S. Lewicki and T. J. Sejnowski, “Learning nonlinear overcomplete representations for efficient coding”, Advances in Neural and Information Processing Systems 10, pp.556-562, 1997.
- 5 M. S. Lewicki and B. A. Olshausen, “A probabilistic framework for the adaptation and comparison of image codes”, J. Opt. Soc. of Am, Vol.16, No.7, pp.1587-1601, 1998.
- 6 T. W. Lee, M. S. Lewicki, M. Girolami, and T. J. Sejnowski, “Blind source separation of more sources than mixtures using overcomplete representations”, IEEE Signal Processing Letters, Vol.6, No.4, pp.87-90, 1999.
- 7 T. Okumura, S. Maekawa, and M. Kotani, “Multichannel blind deconvolution using overcomplete representations, Technical Report of IEICE, NC2002-151, pp.107-112, 2003. (in Japanese)
- 8 N. Yano, S. Maekawa, and T. Yoshimura, “Estimation of single motor unit from the multi-channel surface EMG signals using overcomplete representations”, Technical Report of IEICE, MBE2004-113, pp.5-8, 2005. (in Japanese)
- 9 K. Enokizono, T. Konishi, S. Maekawa, and T. Matushima, “Self-attribution of a viewing object modulates tactile discrimination performance”, 7th International Multisensory Research Forum, Jun. 2006.
- 10 Dipellegrino G, Fadiga L, Fogassi L, Gallese V, and Rizzolatti G, “Understanding motor events - a neurophysiological study”, Exp Brain Res 91: 176-180, 1992.
- 11 Gallese V, Fadiga L, Fogassi L, and Rizzolatti G, “Action recognition in the premotor cortex”, Brain 119: 593-60, 1996.
- 12 Rizzolatti G, Fadiga L, Gallese V, and Fogassi L, “Premotor cortex and the recognition of motor actions”, Cognitive Brain Res 3: 131-141, 1996.

MAEKAWA Satoshi, Ph.D.

*Senior Researcher, Universal City
Group, Knowledge Creating
Communication Research Center
(Former: Senior Researcher, Social
Interaction Group, Keihanna Human
Info-Communication Research Center,
Information and Network Systems
Department)*

*Biological Signal Processing,
Cognitive Psychology, Optics*