On-line Learning Based Electromyogram to Forearm Motion Classifier with Motor Skill Evaluation*

Daisuke NISHIKAWA**, Wenwei YU**,
Masaharu MARUISHI***, Ichiro WATANABE***,
Hiroshi YOKOI**, Yukio MANO***
and Yukinori KAKAZU**

An on-line learning based EMG to motion classifier can manage learning data set by manual appending and automatic elimination compared with conventional off-line learning based classifiers. It is designed to track the alteration of an operator's characteristics through time. The automatic elimination is based on the continuity of human motion. Moreover, in this study we quantify the attainment of motor skill using the classifier. By classifying up to eight forearm motions from two channels of EMG, we investigate the effectiveness of the automatic elimination process, the validity of the attainment of motor skill by seven trials on an unskilled subject, as well as the relationship among the number of electrodes, the classification performance, and the subject's motor skill. Results show that the proposed approaches can simplify decision boundaries, the attainment of motor skill can be used for judging completion of the training by external observers, and bottlenecks in this classifier can be detected.

Key Words: Pattern Recognition, Learning, Biomechanics, EMG to Motion Classifier, On-line Learning, Quantification of Motor Skill

1. Introduction

Surface electromyogram (EMG) is the electric potential detected from the skin surface during muscle contraction. Researchers have employed EMG as a useful source of information or human-machine interfaces. Some promising applications have been developed in the field of rehabilitation engineering, e.g., prosthetic hand/arm⁽¹⁾⁻⁽⁴⁾, and functional electrical stimulation⁽⁵⁾. Further technological innovation and research in this field will enable the handicapped to perform basic ADL (Activities of Daily Living).

Recognizing motion based on EMG signals is one of the obstacles to the development of welfare-related apparatus. Especially, classifying EMG to forearm

motion patterns is more critical than other motions, since the forearm motion classifier must analyze EMG signals with more complex patterns. This is due to the fact that the majority of muscles corresponding to dexterous hand and forearm motions are arranged and overlapped in a limited area. Moreover, the forearm motion classifier must adapt to a wide range of individual variations, e.g., motor learning, and physical modification. This is because that hand and forearm motions are more individual-dependent than motions of the upper arms and lower limbs, which causes more complex EMG patterns with greater diversity.

This problem has attracted the interest of many researchers. Previous works have reported that an original analyzing algorithm and a machine learning method for adaptation were introduced to their classifiers, and their effectiveness was shown⁽³⁾⁻⁽⁹⁾. However, one crucial point that the operator's characteristics usually vary through time has not been sufficiently investigated in almost all the previous works. The systems in previous works can only learn the characteristics in an off-line manner, which

* Received 3rd July, 2000

** Laboratory of Autonomous Systems Engineering, Hokkaido University, Kita-13 Nishi-8, Kita-ku, Sapporo 060-8628, Japan. E-mail: nishi@complex.eng.hokudai.ac.jp

*** Laboratory of Rehabilitation and Physical Medicine, Hokkaido University, Kita-15 Nishi-7, Kita-ku, Sapporo 060-8638, Japan

assumes the operator's characteristics to be stable. It is clear that the characteristics concerned can be altered by motor learning and physical modification. Hence, EMG to motion classifier must have the ability to track the alteration; in other words, it must have an on-line learning ability. Particularly, this ability is important for rehabilitation engineering, since the handicapped are likely to be required to modify their characteristics through training, to correct an inappropriate motion habit or to establish a completely new motion pattern.

For this reason, some researchers have proposed some on-line learning methods. Fukuda et al.⁽¹⁰⁾ have indicated the drawback of off-line learning methods and have improved performance using their on-line adaptive learning method. In Fukuda et al.'s method, the system automatically supplements new learning data according to the similarity to past learning data. Hence, the supplementation is irrelevant to the operators' intention. Besides, their method removes the oldest learning data when new learning data are supplemented, whereas it is clear that the oldest learning data are not necessarily redundant for the classification. On the other hand, we have proposed an EMG prosthetic hand controller using an on-line learning method, which receives operator's evaluation directly and generates learning data when necessary^{(1),(2),(12)}. However, our method does not have the mechanism of removing redundant and harmful learning data. Therefore, we reported the performance deterioration caused by this deficiency^{(2),(12)}.

Furthermore, in all the previous work described above, only the learning and adaptation ability for the systems were evaluated, but operators' level of motor skill was rarely discussed. Namely, it is supposed that the operators have high and fixed motor skill through experiments or training. However, motor skill level is strongly individual-dependent; moreover, it generally alters through time. Hence, for evaluating the ability of a one man-machine system adequately, it is desirable that the machine be able to quantify the operator's motor skill using the interactive information from an operator. Given an operator's motor skill, it is easy to decide how many motions he can control. Moreover, the motor skill level can be used as a basic criterion for judging the cause of low performance in the system. For instance, when the skill level is high and the device shows low performance, it is probable that the motion classifier has one or more bottlenecks somewhere. Especially in the rehabilitation area, the quantification of motor skill can be used for judging completion of the training by external observers such as an occupational therapist or a physical therapist.

In this study, we design an on-line learning based EMG to forearm motion classifier which generates learning data according to an operator's intention and eliminates redundant or harmful learning data. Moreover, we propose a method for quantifying attained motor skills. Experiments show the effectiveness of these approaches.

This paper contains five sections. In section 2, an on-line learning based EMG to forearm motion classifier will be described. In section 3, an algorithm for quantifying motor skill attained will be made clear in detail. Experiments were performed to verify the effectiveness of our classifier and quantification method, of which the details will be explained in section 4. Finally, a brief conclusion will be given.

2. Design of On-line Learning Based EMG to Forearm Motion Classifier

2.1 Outline

Based on the general methodology of pattern recognition, the conventional EMG to forearm motion classifier employs two processes. One is a feature extraction process, which generates a feature vector from EMG signals. The another is a classification process from the feature vector to a forearm motion. Moreover, the classifier must obtain adaptively the relationship between EMG and forearm motion so that the individual variation can be dealt with. Hence, the classifier also employs a learning mechanism embedded in the classification process. In this paper, we call the unit taking charge of feature extraction the *analysis unit*, and the unit taking charge of classification the *adaptation unit*.

In addition, for tracking the alteration of individual variation through time, the proposed classifier should contain a mechanism supervising a learning data set, in which learning data are managed by appending and elimination processes. These two processes should not depend on the learning mechanism in the classification process. Therefore, we design a novel unit for supervision, and term it *trainer unit*.

The schematic diagram of the classifier is shown in the Fig. 1. Conventional approaches perform only feature extraction and the classification, corresponding to the analysis unit and the adaptation unit in this figure, respectively. On the other hand, our proposed on-line learning method supervises the learning data set by the trainer unit.

Our previous studies^{(1),(2),(12)} have shown the effectiveness of the analysis and adaptation methods employed in this paper. Hence, we omit the detailed description of these methods.

Fig. 1 Schematic diagram of the on-line EMG to forearm motion classifier

2.2 Analysis unit

The EMG is complex signal and some analysis methods are needed to extract motor information. Hence, many researchers have proposed EMG signal analysis methods, such as mean absolute value^{(4),(10)}, frequency analyses^{(6),(7),(9)}, and heuristic approaches^{(3),(8)}. Besides, some researchers have discussed the differences between these analysis methods in terms of their effectiveness to classify motions^{(3),(9)}. They have concluded that suitable analysis methods enable systems to classify several motions from fewer EMG signals.

Our analysis unit generates the feature vector $\mathbf{p}$ from EMG signals $\mathbf{s}$. Based on our experience in previous research, we employ the mean absolute value and the Gabor transform^{(2),(12)} for analyzing algorithm:

$$w(\mathbf{s}) = \mathbf{p}. \quad (1)$$

Here, w stands for the function of feature extraction. More details and relevance are discussed in Refs. (1), (2), and (12).

2.3 Adaptation unit

The mapping relationship between feature vectors and motion is probably nonlinear in prospect. Moreover, operator's characteristics differ with different people. Therefore, to classify many motions accurately, an adaptation mechanism is necessary to adjust the mapping function to absorb the individual variation.

In previous works, some adaptation mechanisms, such as artificial neural networks⁽⁶⁾⁻⁽¹⁰⁾, k -nearest neighbor⁽³⁾, evolvable hardware⁽⁴⁾, and linear discriminant analysis⁽⁹⁾, have been employed. Some achievements have been reported: devices that can discriminate six forearm motions from four channels of EMG⁽¹⁰⁾, six forearm motions from two channels of EMG^{(4),(7)}, four upper arm motions from two channels of EMG^{(3),(8),(9)}, as well as three thumb motions from one channel of EMG⁽⁶⁾, were developed. These results were obviously influenced by not only the adaptation mechanisms but also the analysis methods. Therefore, we cannot conclude clearly what is the most superior approach among all previous works.

Our adaptation unit uses a three-layered feed-forward neural network (NN) for the nonlinear mapping function^{(1),(2),(12)}. From a feature vector $\mathbf{p}$, the adaptation unit calculates an output $\mathbf{o}$ in the NN section.

$$\mathbf{o} = m(\mathbf{p}). \quad (2)$$

Here, m stands for the NN mapping function. More details and relevance are discussed in Refs. (1), (2), and (12).

2.4 Trainer unit

In our previous studies, we determined the essential function of on-line learning. The function is to realize the effective learning of an operator's characteristics from a few pieces of learning data provided occasionally by the operator^{(1),(2),(12)}. In addition to the function, in this paper we indicate another required function, which is the elimination of redundant or harmful learning data. It prevents complicating the clusters and their relations while maintaining the total performance. These functions are described below.

The trainer unit generates the learning data that contain a teaching signal from an operator and the feature vector from analysis unit at the moment. When the trainer unit receives the teaching signal, it composes the learning data. Then, the obtained new learning data are added to the learning data set. Moreover, the trainer unit updates the inner states of the adaptation unit in the interval of main processing, which involves classifying the forearm motion from EMG signal. In the updating process, the trainer unit sends the learning data to the adaptation unit in order, receives its output, and calculates a square error value from the output and the learning data. When the trainer unit judges that the output is not correct compared with the learning data or the square error is above a threshold value, it starts the learning of the adaptation unit. The trainer unit employs the back-propagation algorithm as the learning method. To shorten the learning steps, the trainer unit employs a kind of active learning method, which alters the learning rate by the average of the square error through all the learning data sets^{(2),(12)}.

To choose surplus or harmful learning data for the elimination process, the trainer unit should contain a suitable valuation standard. In this paper, we employ and investigate three candidates for the performance functions to calculate the degree of contribution to the classification by certain learning data. Based on the calculated performance, a learning data elimination procedure will be executed.

At the beginning, an expression of the performance function, i.e., the degree of contribution, is defined as:

$$\varphi = \langle \mathbf{p}, \theta, a \rangle. \quad (3)$$

$$a_{t+1} \equiv R \cdot a_t + f, \quad (4)$$

$$0 < R < 1, \quad (5)$$

$$0 \leq f \leq 1. \quad (6)$$

In these formulae, φ is learning datum; θ is a class ID in the $\Theta = \{\theta_1, \theta_2, \dots, \theta_{|\Theta|}\}$ in which $|\Theta|$ means the total number of classes; a is the degree of contribution; R is a damping factor to gradually decrease the degree of contribution through time; and f is one of the performance functions defined in the following section. When a of one learning datum is relatively larger than that of other learning data, the learning datum φ is more important in that situation.

The learning data with low degree of contribution are of little importance for the classification. Hence, a learning data elimination procedure will be triggered to handle the learning data. The process is shown as follows.

$$\begin{aligned} \Phi^\omega &\equiv \{\varphi_1, \varphi_2, \dots, \varphi_{|\Phi^\omega|}\} \\ &= \{\langle \mathbf{p}_1, \omega, a_1 \rangle, \langle \mathbf{p}_2, \omega, a_2 \rangle, \dots, \langle \mathbf{p}_{|\Phi^\omega|}, \omega, a_{|\Phi^\omega|} \rangle\}. \end{aligned} \quad (7)$$

$$a_\varphi / a_{\max} < B \Rightarrow \Phi^\omega = \Phi^\omega - \{\varphi\}, \quad (8)$$

$$a_{\max} \equiv \max_{\Phi^\omega}(a), \quad (9)$$

$$0 \leq B < 1. \quad (10)$$

In these formulae, B is the threshold determining learning data elimination; and Φ^ω is the learning data set belonging to the same class ω . In the case of $B=0$, the learning data elimination procedure is not executed. This corresponds to the system in our previous works^{(1),(2),(12)}. a represents the contribution value of currently evaluated learning data and $a_{\max}$ represents the highest contribution value in the identical class ω . Formulae (7)-(10) mean that if the ratio of a to $a_{\max}$ is lower than a constant threshold B , then the currently evaluated learning data will be eliminated.

In this paper, we define three candidates for performance functions $\{f_1, f_2, f_3\}$. These functions must have a property that values of functions are in direct proportion to the importance of the learning data. Besides, these functions are likely to be problem-dependent; hence, we employ heuristic functions in this study.

The performance function f is defined on the basis of the following hypotheses:

(1) When the trainer unit observes that the sequence of the classification results in a fixed period of time and all the results belong to one same class, it decides the classification at that time to be successful. Here, even if our system consecutively classifies EMG as the identical and incorrect class, it is immediately rectified by the operator's teaching. This hypothesis becomes existent on the premise that there exists a certain period during which humans cannot switch one

motion to another.

(2) When the trainer unit decides the classification to be successful and receives a feature vector $\mathbf{p}$ and a classification result ω , it takes learning datum φ in Φ^ω , which has the closest distance to $\mathbf{p}$, as most important one for this classification. In the case that the adaptation unit employs nearest neighbor or other similar algorithms as the mapping function, this hypothesis is within reason.

From hypotheses (1) and (2), the performance functions are defined. When a series of outputs Ψ , which contain the feature vectors and the classification results for the past Q steps at time s , are given,

$$\Psi(s) \equiv \{\psi_{s-Q}, \psi_{s-Q-1}, \dots, \psi_{s-1}, \psi_s\}, \quad (11)$$

$$\psi \equiv \langle \mathbf{p}, \theta \rangle. \quad (12)$$

Moreover, $\theta_t = \omega\{t=s-Q, \dots, s\}$ holds. Then, we define a fundamental performance function $h_\varphi(\psi)$ which is used for actual performance functions as:

$$h_\varphi(\psi) \equiv \begin{cases} 1 & : \varphi = \arg \min_{\varphi \in \Phi^\omega} \{g(\varphi, \psi)\}, \\ 0 & : \text{otherwise.} \end{cases} \quad (13)$$

$$g(\varphi, \psi) \equiv \|\mathbf{p}_\varphi - \mathbf{p}_\psi\|. \quad (14)$$

This function means that the existing learning datum nearest to the feature vector from the analysis unit gets a reward and others get no the reward. To apply this function to a series of outputs $\Psi(s)$, we consider three candidates for performance function:

$$f_1 \equiv \max_{t=s-Q, \dots, s} \{h(\psi_t)\}. \quad (15)$$

$$f_2 \equiv \frac{1}{Q} \sum_{t=s-Q}^s h(\psi_t). \quad (16)$$

$$f_3 \equiv -R \cdot a_t + f_1. \quad (17)$$

Formula (15) means that a group of learning data which contribute to the classification only once or more within the period will get the same amount of reward impartially. On the contrary, formula (16) means that the more a learning datum contributes, the more it gets a reward. Therefore, this method is more 'fruition-ism' than formula (15). Besides, formula (17) differs from formula (15) in that it discounts the past reward when the learning datum gets a reward. Therefore, it is the most 'egalitarianism' in the other methods.

Experiments will investigate the difference among the three candidates.

3. Quantifying Cluster Complexity and Motor Skill

To design a method for classifying human motion, we consider the operator's repeatability, which closely relates to the motor skill⁽¹³⁾. In our on-line system, the low repeatability of motor program unavoidably causes the low separability of the EMG signal. Moreover, the repeatability increases as the skill forms and improves, so that the estimation of the

repeatability makes it possible to quantify the motor skill through operator's training, which is particularly important in the rehabilitation area. Bootsma et al.⁽¹⁴⁾ used another term 'consistency' to claim almost the same thought.

In order to evaluate the repeatability, we employ the complexity of clusters formed by EMG feature vectors in the learning data set. Further, a Davies's cluster separation measure⁽¹¹⁾ in the two-dimensional Euclidean space is used to calculate the complexity standard. Then, we term the measure $\bar{d}$ as *cluster complexity*, which is formulated as:

$$\bar{d} \equiv \frac{1}{|\Theta|} \sum_{i=1, \dots, |\Theta|} \max_{j=1, \dots, i-1, i+1, \dots, |\Theta|} \{d_{ij}\}. \quad (18)$$

$$d_{ij} \equiv \frac{u_i + u_j}{\|\mathbf{v}_i - \mathbf{v}_j\|}. \quad (19)$$

$$u_i \equiv \sqrt{\frac{1}{|\Phi^{\theta_i}|} \sum_{k=1}^{|\Phi^{\theta_i}|} \|\mathbf{p}_k - \mathbf{v}_i\|^2}. \quad (20)$$

In these formulae, $\mathbf{v}_i$ stands for the centroid of feature vectors in class θ_i ; u_i is the standard deviation in the same class θ_i ; and $|\Phi^{\theta_i}|$ means the total number of elements in a learning data set Φ^{θ_i} . From formulae (1) and (7), $\mathbf{p}$ is the feature vector in element φ in the learning data set Φ . When measure $\bar{d}$ becomes low/high, the classes become more separated/intertwined with each other.

In the initial stage of training, the clusters are widely dispersed because of the low repeatability. Hence, the cluster complexity $\bar{d}$ takes a high value. In contrast, at the end of the training, the clusters are adequately concentrated. Hence, the cluster complexity is relatively low. Therefore, comparing the value of cluster complexity through the training, we can monitor the change in operators' motor skill. However, using our approach, the completion of the training cannot be determined because we cannot find a desired value to be compared with. Moreover, because of individual variations, it is also difficult to find an individual-independent criterion to denote the completion of the training.

To investigate the attainment of motor skill for different people, we use two types of cluster complexities. One is applied to the entire learning data ($\bar{d}_{all}$), and the other is applied to the learning data excluding the eliminated learning data ($\bar{d}_{survivor}$). As the training proceeds and nearly reaches completion, the repeatability is high, so the opportunity for the elimination process is low except for the redundant learning data. Therefore, the difference between these two parameters decreases ($\bar{d}_{survivor}/\bar{d}_{all} \approx 1$). On the contrary, the low repeatability leads to the high opportunity for the elimination process. Besides, $\bar{d}_{all}$ is generally higher than $\bar{d}_{survivor}$ in this situation. This is because the eliminated learning data usually exist in

Fig. 2 Experimental setup

the outlying region of a cluster. Hence, the difference increases and their quotient tends to zero ($\bar{d}_{survivor}/\bar{d}_{all} \approx 0$). We define the *attainment of motor skill* as the quotient $\bar{d}_{survivor}/\bar{d}_{all}$.

The reason for the rejection of the strategy of simply comparing a number of eliminated learning data is that this number strongly depends on the number of total teaching and it weakly depends on the complexity of clusters.

4. Experiments

We apply the proposed classifier to an EMG prosthetic hand controller. Through the experiments, we clarify the effectiveness and necessity of the proposed approaches. Namely, following points should be clarified: The elimination method makes it possible to simplify the decision boundaries; hence, it facilitates the classification. The combination of elimination method and motor skill quantification realizes the judgment of training completion. Moreover, it gives us the possibility to detect the bottleneck of the man-machine system such as an EMG prosthetic hand.

4.1 Setup

We want to show that our proposed system is effective despite differences between normal people and amputees. Hence, one skilled normal subject (Subj. A), one unskilled normal subject (Subj. B), and one skilled amputee who suffers from wrist disarticulation (Subj. C) take part in the experiments. Subj. A and Subj. B are in their twenties. Subj. C is 50 years old.

Figure 2 shows the experimental setup. The proposed EMG classifier is implemented by software in PC. For sending the teaching signal to the trainer unit, a keyboard is used. In this study, a CG of a human hand is used as a substitute of a real prosthetic hand. The subjects are required to control the CG hand, and they send a teaching signal when they judge the CG hand as moving unsatisfactorily. The teaching signal is class ID corresponding to the motion performed by the operator at that time. Based on

Table 1 Common parameters in the experiments

<i>analysis unit</i>			
sampling rate	1600 Hz	number of electrodes	2 or 3
<i>adaptation unit</i>		<i>trainer unit</i>	
number of units in middle layer	18	Q	20

Fig. 3 Eight forearm motions to be classified in the experiment

pioneering studies, one teaching signal generates five learning data.

Dry-type electrodes are used and placed on the skin surface near the elbow. This is because that area is the origin of many forearm muscles, so that the detected EMG signals are synthesized and contain information of multiple motions. Another advantage is that we can measure the EMG at the same place even for severer upper types of forearm amputation⁽¹²⁾.

In this study, our system attempts to classify up to eight forearm motions, including two forearm motions, two wrist motions, as well as four hand motions. Each motion is assigned an ID number (Fig. 3). One reason for separating the thumb from the other fingers is that the thumb opposition plays an important role in daily hand function^{(2),(12)}.

Table 1 shows the common parameters used in the experiments. These values were obtained from our previous works^{(1),(2),(12)} and pioneering studies. The sampling rate is set based on the bandwidth of the amplifier in the electrode, which is 20 ± 5 Hz to 450 ± 50 Hz. The response time is 10~15 milliseconds and contains the process of classifying EMG to forearm motion and the process of learning the subject's characteristics. The response time is adequately short for controlling a prosthetic hand.

For quantifying the classification ability, subjects

are required to take an ability test⁽¹⁾. In the ability test, the system instructs the subjects to execute the motion shown in the monitor and it compares the output with the instructed motion at that time. The system calculates the sequence of the output for three seconds per one instruction. Depending on the response time, the number of outputs for one instruction is about 200~300. Besides, each motion is instructed twice in the test. We use a *classification rate* as the index of comparison.

4.2 Relevance of the elimination process to the performance function (experiment 1)

In the beginning, we discuss the relevance of the elimination process to the performance function which is one of the three candidates (15)~(17), as described in section 2.4. For this purpose, we will clarify the differences in characteristics among these candidates by investigating the classification rate and the cluster complexities.

In the elimination method (3)~(17), the number of parameters is three (R, B, Q). Here, step Q depends on the problem setting. This is because the proposed classifier conforms to the presupposition that humans cannot switch one motion to another within a certain period of time. Moreover, the presupposition does not depend on the performance function. Hence, the parameter Q is fixed at 20 (Table 1), which means that the period is defined as about 200~300 milliseconds from the response time.

One skilled normal subject (Subj. A) takes part in this experiment. Many trials are required for searching parameters adequately. Hence, it is reasonable to select a skilled subject who rarely changes his motor skill through the trials. However, considering the heavy workload of many trials, we did not include the amputee subject in this experiment. For each pair of parameters R and B , two trials are performed. Eventually, the total number of trials is 58. For each trial, the classification rate and the cluster complexities ($\bar{d}_{survivor}$, $\bar{d}_{all}$) are measured. The number of motions to be classified is eight (ID01~ID08). The number of electrodes is two.

Figure 4 shows the results. The solid lines stand for the classification rate described above; the dashed lines stand for the cluster complexity of ($\bar{d}_{survivor}$), that is the learning data set excluding eliminated learning data. In each graph, mean values for two trials are plotted. In all cases, the classification rate and the cluster complexity present decreasing trend with increasing threshold B . However, for all the three candidates, the decrease is clearly different from each other. This indicates that these three candidates for performance function differently and strongly affect learning data elimination; hence, they should

Fig. 4 Differences in characteristics of three candidates (experiment 1)

also affect the decision boundaries and the clusters' mutual relations in a different way.

Besides, in the case of no-elimination ($B=0.00$), the mean classification rate in the four trials is 93.0% and the cluster complexity $\bar{d}_{all}$ is 0.150. Because by some parameter combinations, the cluster complexities are nearly even to that of the no-elimination case, the eliminating algorithms with such parameters have no effect on cluster simplification. As mentioned before, the purpose of the elimination method is to prevent cluster complexity while maintaining its performance. From this viewpoint, in the case of performance function f_2 , the elimination method is less effective with a relatively large damping factor R . Meanwhile, the elimination method exerts a bad influence on the classification rate with relatively small R . In the same way, the candidate f_3 is less effective for cluster simplification.

Consequently, only the performance function f_1 meets the requirement. In the following experiments, f_1 is employed as the performance function, with parameters R and B being 0.90 and 0.025, respectively.

Naturally, more appropriate performance functions may exist. Particularly, in this study, the classifier only uses the successful classification to calculate performance functions although it is also reasonable to use the failing classification. Hence, one of our future works will be the improvement of performance function that can deal with both reward and penalty.

4.3 Validity of attainment of motor skill (experiment 2)

Next, we verify the validity of the attainment of motor skill proposed in section 3. For discussing the

Fig. 5 Process of training for Subj. B (experiment 2)

attainment of motor skill, the unskilled subject would play an essential role. Hence, one unskilled subject (Subj. B) is required to participate in the experiment. Unfortunately, we cannot employ the unskilled amputee as subject to make a complete investigation. The classification rate, the cluster complexities, as well as the attainment of motor skill in each experiment are measured and calculated. The number of motions to be classified is eight (ID01~ID08). The number of electrodes is two. The subject is required to conduct the experiment once every day and altogether seven trials are conducted.

Figure 5 shows the process of training. (a) shows the relationship between the classification rate and the attainment of motor skill; (b) gives two types of cluster complexities of the learning data set. Through the trials, the classification rate increases, and finally reaches 90%. This means that the subject is acquiring the motor skill through the trials. Moreover, the attainment increases with the classification rate, and finally reaches over 0.8 point. This indicates that the attainment meets the demand for a valid index detecting operator's motor skill. Besides, compared to $\bar{d}_{all}$ which decreases through the trials, $\bar{d}_{survivor}$ barely varies. We infer that the elimination method restrains the capacity of class; hence, it restricts the variation of $\bar{d}_{survivor}$.

In order to confirm the phenomenon of cluster simplification caused by the elimination method and the operator's training, we plot feature vectors in the learning data set on a two-dimensional distribution map (Fig. 6). Upper maps show the initial stage of training (2nd trial). Lower maps show the end stage of training (6th trial). The maps on the right are the results when the eliminated learning datum is excluded. The maps on the left are the results with the eliminated learning data. For drawing intelligible maps, we quasi-optimize the parameters of linear transformation and normalize the X and Y axes. Here, the projection is made from a multi-dimensional feature vector to the two-dimensional space

Fig. 6 Feature vectors plotted in two-dimensional distribution map (experiment 2)

using a linear transformation. Hence, the overlapping of clusters do not always indicate impossibility of classification but it means difficulty of classification. In both trials, the elimination method simplifies the clusters and their relationships. However, in the maps with eliminated learning data (a, c), there are some differences on the decision boundaries and the cluster relations. That is, in the initial stage of training (a), distributions of clusters are large, and especially, such distributions are noticeable in classes ID07 and ID03. Moreover, the extent of the overlapping area is also larger in the initial stage of training (a) than in the end stage of training, e.g., the lump including ID02, ID07 and ID06. These differences cause the gap of the cluster complexities and the classification rates between the initial stage and the end stage of training. It is clear that the training lessens the overlapping of the clusters.

Besides, even if the elimination method simplifies the clusters and their relations in the initial stage of the training, it does not lead to a high classification rate (Fig. 5). We infer the cause of this phenomenon in the following: In the initial stage, motion repeatability is too low because the cluster capacity is smaller than that required to express the dispersion of learning data. Namely, in the initial stage of training, the generated clusters only represent a part of real clusters. Obviously, the incomplete clusters cause the misclassification.

In our current EMG analysis method, even if the classifier employs a sufficiently large cluster capacity, it cannot avoid the overlapping of clusters, and it gets

a low classification rate. We have verified it in the experiment on unskilled persons through the no-elimination method^{(2),(12)}. Therefore, for improving the classification ability for unskilled people, we will study the more flexible elimination method and the more effective EMG analysis method concurrently.

Moreover, the proposed elimination method can only compare the learning data in the same class. It cannot evaluate the relationship between learning data of different classes. However, the inter-cluster relationship plays one of the important roles. Hence, one of our future works will be the development of an evaluation method of inter-cluster relationships.

4.4 Probability of detecting the bottleneck in the man-machine system (experiment 3)

Finally, to discuss the bottleneck in the man-machine system such as the EMG prosthetic hand, we investigate the relationship between the classification rate and the attainment of motor skill. To discuss it in detail, an identical subject is required to conduct an experiment using various combinations of parameters including the number of motions and electrodes.

One skilled amputee subject takes part in this experiment. She has had experience in our previous experiments⁽¹²⁾ and is able to control four forearm and wrist motions (ID01~ID04) and four hand motions (ID05~ID08) separately. However, she could not control all the eight motions. We infer from previous studies that it is the information insufficiency of detected EMG signals that caused this phenomenon.

The experimental procedure is as follows: To begin with, the subject is required to control four motions (ID01~ID04) with two electrodes. If she achieves motion control with high accuracy, which is judged by the classification rate, she is required to control two more motions. However, if she cannot achieve this, she is required to control the same number of motions using one more electrode. For each pair of parameters, several trials are performed.

Table 2 shows the results. In spite of various settings, all the attainments of motor skill are high and nearly constant. This indicates that the subject is at the stage of training completion. Therefore, the classification rate indicates the ability of the classifier. From the results, the addition of electrodes may increase the ability of information extraction in this experiment. To clarify this, Fig. 7 shows the differences in clusters and their relationships concerning the number of electrodes. These maps indicate that the increase in the number of electrodes results in the gathering within each cluster. Moreover, it simplifies the relationship between the clusters. Here, in the case that the number of motions is eight, the subject received a low classification rate (Table 2),

Table 2 Relationship between experimental settings and results for the same subject (experiment 3)

number of motions	L	number of trials	classification rate (%)	attainment of motor skill
4	2	7	90.1	0.95
6	2	3	75.1	0.93
6	3	3	88.5	0.93
8	3	3	73.5	0.95

Fig. 7 Improvement of extracted information by adding electrode (experiment 3)

though she got a high attainment of motor skill. The discussion suggests that the classification rate can be enhanced further by improvement of information extraction.

Table 3 shows the results of all subjects in classifying eight motions. Subj. A got a high classification rate and a high attainment of motor skill; hence, he is in the stage of training completion and needs less improvement for the classifier. Subj. B in the first trial got low attainment of motor skill. This indicates that she is in the starting stage of training, and must undergo training to improve her performance. In the last trial, she got a high classification rate but her attainment of motor skill is still inadequate. Therefore, more training may improve her classification rate. On the other hand, Subj. C got a high attainment of motor skill but got a relatively low classification rate. This indicates that she is in the stage of training completion and the low performance is caused by the ability of the classifier. Hence, some bottlenecks exist somewhere, e.g., the information extraction.

Although Subj. C can control six or eight motions by increasing the number of electrodes, Subj. A and Subj. B achieved a high classification rate using only two channels of electrodes. We infer that the difference in the absolute levels of motor skill determines the necessary number of electrodes. Accordingly, Fig. 8 shows the differences in clusters and their relationships in other two subjects. In spite of using

Table 3 Detecting the bottleneck of the system (experiment 3)

subject	L	trial	classification rate (%)	attainment of motor skill
Subj. A	2	2nd	95.8	0.96
Subj. B	2	1st	59.3	0.33
Subj. B	2	7th	90.7	0.80
Subj. C	3	2nd	74.7	0.98

Fig. 8 Differences of clusters and their relations in two subjects (experiment 3)

fewer electrodes, each cluster about Subj. A is more centralized whereas the clusters are separated from each other. This indicates that even if the attainment of motor skill is nearly even, the absolute level of motor skill differs for each person. This difference may be caused by age. As a matter of fact, Subj. C is about thirty years older than the other subjects. Consequently, one of our future works will be to quantify the absolute level of motor skill for more effective training and estimation of the training.

5. Conclusions

In this study, we designed an on-line learning based EMG to forearm motion classifier which generates learning data according to the operator's intention and eliminates redundant or harmful learning data. Moreover, we proposed a quantification method for the attainment of motor skill. Three kinds of experiments showed the effectiveness of the proposed elimination and quantification method. We also discussed the characteristics of the proposed three types of performance functions, the validity of the quantification of the attainment of motor skill, as well as the probability of detecting the bottleneck in the man-machine system by comparing the classification rate and the attainment of motor skill.

In comparison with previous on-line learning approaches^{(1),(2),(10),(12)}, we showed that the proposed EMG to forearm motion classifier has the advantages

of maintaining the clusters simply, detecting the progress of the operator's motor learning, as well as clarifying the bottleneck of the system. These abilities are effective not only in difficult-to-predict and nonstationary systems including most man-machine systems, but also in the area requiring quantitative estimation of a person's dexterity, such as rehabilitation engineering.

The future directions of this study will include the improvement of performance function that can deal with both reward and penalty, the development of an evaluation method of inter-cluster relationships, the innovation of EMG analysis methods, as well as the quantification of the absolute level of motor skill.

This research is supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture, Japan.

References

- (1) Nishikawa, D., Yu, W., Yokoi, H. and Kakazu, Y., On-Line Learning Method for EMG Prosthetic Hand Controlling, Trans. of the Institute of Electronics, Information and Communication Engineers D-II, (in Japanese), Vol. J82-DII, No. 9 (1999), pp. 1510-1519.
- (2) Nishikawa, D., Yu, W., Yokoi, H. and Kakazu, Y., EMG Prosthetic Hand Controller Using Real-time Learning Method, Proc. of the 1999 IEEE Systems, Man, and Cybernetics Conf., (1999), pp. I153-I158.
- (3) Zardoshti-Kermani, M., Wheeler, B.C., Badie, K. and Hashemi, R.M., EMG Feature Evaluation for Movement Control of Upper Extremity Prostheses, IEEE Trans. on Rehabilitation Eng., Vol. 3, No. 4 (1995), pp. 324-333.
- (4) Kajitani, I., Hoshino, T., Kajihara, N., Iwata, M. and Higuchi, T., An Evolvable Hardware Chip and Its Application as a Multi-Function Prosthetic Hand Controller, Proc. of the 16th National Conf. on Artificial Intelligence, (1999), pp. 182-187.
- (5) Handa, H., Takahashi, H., Saito, C., Ichie, M., Handa, Y., Kameyama, J. and Hosimiya, N., Development of an FES System Controlled by EMG Signals, Ann. Int. Conf. of the IEEE Eng. in Medicine and Biology Society, Vol. 12, No. 5 (1990), pp. 2773-2774.
- (6) Farry, K.A., Walker, I.D. and Baraniuk, R.G., Myoelectric Teleoperation of a Complex Robotic Hand, IEEE Trans. Robotic and Automation, Vol. 12, No. 5 (1996), pp. 775-788.
- (7) Uchida, M., Ide, H. and Ninomija, S.P., Control of a Robot Arm by Myoelectric Potential, Journal of Robotics and Mechatronics, Vol. 5, No. 3 (1993), pp. 259-265.
- (8) Hudgins, B., Parker, P. and Scotto, R.N., A New Strategy for Multifunction Myoelectric Control, IEEE Trans. Biomedical Eng., Vol. 40, No. 1 (1993), pp. 82-94.
- (9) Englehart, K., Hudgins, B., Parker, P.A. and Stevenson, M., Classification of the Myoelectric Signal using Time-Frequency Based Representations, Medical Eng. and Physics, Vol. 21 (1999), pp. 431-438.
- (10) Fukuda, O., Tsuji, T. and Kaneko, M., An EMG Controlled Robotic Manipulator Using Neural Networks, Proc. of IEEE Int. Workshop on Robot and Human Communication, (1997), pp. 442-447.
- (11) Davis, D.L. and Bouldin, D.W., A Cluster Separation Measure, IEEE Tran. on Pattern Analysis and Machine Intelligence, Vol. PAMI-1, No. 2 (1979), pp. 224-227.
- (12) Nishikawa, D., Yamaguchi, H., Yu, W., Maruishi, M., Watanabe, I., Yokoi, H., Yukio, M. and Kakazu, Y., Discriminating Ten Forearm Motions by EMG Prosthetic Hand Controller Using Real-Time Learning Method, (in Japanese), Proc. of Biomechanism Symposium, (1999), pp. 201-212.
- (13) Fumoto, N., Verification of Learning Effect in Physical Exercise, J. of Health, Physical Education and Recreation, (in Japanese), Vol. 39 (1989), pp. 598-605.
- (14) Bootsma, R.J. and Van Wieringen, P.C.W., Timing an Attacking Forehand Drive in Table Tennis, J. of Experimental Psychology: Human Perception and Performance, Vol. 16, No. 1 (1990), pp. 21-29.