A Class Selection Method Based on a Partial Kullback-Leibler Information Measure for Biological Signal Classification

Taro SHIBANOKI, Keisuke SHIMA, Takeshi TAKAKI, Toshio TSUJI,
Akira OTSUKA and Takaaki CHIN

Abstract— This paper proposes a novel class selection method based on the Kullback-Leibler (KL) information measure and outlines its application to optimal motion selection for bioelectric signal classification. When a user has no experience of controlling devices using bioelectric signals, for instance controlling a prosthetic hand using EMG signals, it is well known that voluntary generation of such signals might be difficult, so that the classification issue of multiple motions thus becomes problematic as the number of motions increases. An effective selection method for motions (classes) is needed for accurate classification. In the proposed method, the probability density functions (pdfs) of measured data are estimated through learning involving a multidimensional probabilistic neural network (PNN) based on the KL information theory. A partial KL information measure is then defined to evaluate the contribution of each class for classification. Effective classes can be selected by eliminating ineffective ones based on the partial KL information one by one.

In the experiments performed, the proposed method was applied to motion selection with three subjects, and effective classes were selected from all motions measured in advance. The average classification rate using selected motions under the proposed method was 92.5 ± 0.9 %. These outcomes indicate that the proposed method can be used to select effective motions for accurate classification.

I. INTRODUCTION

Biological signals such as electromyograms (EMGs) and electroencephalograms (EEGs) reflect human intentions and motions, and can be used to control a range of machines that estimate these variables. Various methods to classify biological signal patterns have been proposed using the AR model [1], discriminant analysis [2], neural networks (NNs) [3], [4] and other approaches. Tsuji *et al.* proposed a log-linearized Gaussian mixture network (LLGMN) [5] that includes a Gaussian mixture model (GMM), and confirmed its effectiveness in biological signal classification [6], [7]. However, if users cannot stably generate biological signals, the difference between motions (classes) becomes ambiguous and accurate classification is difficult.

Generally, pattern classification is performed by estimating the relationship between learning data (biological signals) and class labels (the corresponding intentions). Cluster analysis (in which a collection of patterns is organized into clusters based on similarity) is often adopted to estimate the number of classes based on class structure. This type of analysis can be dichotomized as either hierarchical clustering [8] or non-hierarchical clustring (with the k-means algorithm) [9].

Hierarchical clustering involves repeated hierarchical grouping of subclasses and a final grouping with a suitable number of classes. The k-means algorithm, which is simple and has local optimality, identifies a partition of the input space with the number of classes defined in advance. In these methods, however, it is difficult to achieve appropriate clustering for data whose distribution is unknown.

This paper proposes a class selection method based on Kullback-Leibler (KL) information and its application to motion selection for EMG classification. The authors previously proposed a partial KL information measure as a metric regarding effective dimensions of input vector for classification, and indicated its effectiveness for selecting optimal channel positions [10]. In this paper, we consider application of the partial KL information measure as a metric for evaluating effective classes. In the proposed method, measured signals are regarded as probability variables, and their probability density functions (pdfs) are estimated using probabilistic neural network (PNN) learning based on KL information. In addition, classes that are difficult to classify accurately are eliminated one by one based on a partial KL information measure. The proposed method facilitates resolution of the classification problem.

Section II introduces the proposed class selection method and a partial KL information measure, Section III describes the EMG classification method for applying the proposed technique to motion selection, and Section IV reports on class (motion) selection ability evaluation through simulation experiments and EMG classification experiments.

II. CLASS SELECTION METHOD BASED ON A PARTIAL KL INFORMATION MEASURE

A novel class selection method based on a partial KL information measure is described in this section.

A. Partial KL Information Measure

A common way of classifying non-Gaussian and/or non-linear data is to regard them as probability variables and estimate their pdfs. In this way, accurate estimation of data distribution enables accurate classification of data. The KL information represents the difference between two probability distributions [13], and nears zero as the estimated distribution comes closer to the true distribution. It is possible to

T. SHIBANOKI, K. SHIMA, T. TSUJI and T. TAKAKI are with Graduate school of Engineering, Hiroshima University, 1-4-1 Kagamiyama Higashi-Hiroshima Hiroshima, JAPAN shibanoki@bsys.hiroshima-u.ac.jp

A. OTSUKA is with Faculty of Welfare and Healthcare, Prefectural University of Hiroshima, 1-1 Gakuen-cho Mihara Hiroshima, JAPAN

T. CHIN is Hyogo Rehabilitation Center 1070 Akebono-cho Nishi-ku Kobe Hyogo, JAPAN

evaluate classes (for classification or not) using the estimated distribution. This section describes KL information and the proposed partial KL information.

Let us consider a case that N_k samples belong to the class k $(k = 1, 2, \cdots, K)$ and each sample is an L-dimension variable vector $(\boldsymbol{x} = [x_1, x_2, \cdots, x_L] \in \Re^L)$. It is assumed that $\boldsymbol{x}_n^{(k)} \in \Re^L$ is the probability variable vector of the nth sample in the kth class, and that the probabilities of the kth class in the true probability distribution and the estimated distribution are P_n^k and Q_n^k, respectively. The KL information is calculated using

$$
\begin{aligned}
I_n(\boldsymbol{P}, \boldsymbol{Q}) &= \sum_{k=1}^{K} P_n^k \log \frac{P_n^k}{Q_n^k} \\
&= \sum_{k=1}^{K} P_n^k \log P_n^k - \sum_{k=1}^{K} P_n^k \log Q_n^k, \quad (1)
\end{aligned}
$$

where $\boldsymbol{P} = [P_n^1, P_n^2, \cdots, P_n^K]^{\mathrm{T}}$ and $\boldsymbol{Q} = [Q_n^1, Q_n^2, \cdots, Q_n^K]^{\mathrm{T}}$, and the first term on the right side is reduced to a constant; accordingly, if the second term reaches a minimum, then so does the KL information. This means that the estimated distribution is close to the true distribution, and if the two correspond, the KL information becomes zero.

The authors proposed a partial KL information measure as a metric of variable selection [10]. When applying the partial KL information as a metric for evaluating effective classes, since the number of data changes when a class is eliminated, it is difficult to evaluate the accuracy of data distribution estimates on the same scale using such information. Against such a background, this paper newly defines partial KL information to evaluate classes:

$$
E_{[i']} = \frac{I(\boldsymbol{P}, \boldsymbol{Q})}{I(\boldsymbol{P}_{[i']}, \boldsymbol{Q}_{[i']})} = \frac{N}{N'} \frac{\sum_{n=1}^{N} I_n(\boldsymbol{P}, \boldsymbol{Q})}{\sum_{n=1}^{N'} I_n(\boldsymbol{P}_{[i']}, \boldsymbol{Q}_{[i']})}, \quad (2)
$$

$I_n(\boldsymbol{P}_{[i']}, \boldsymbol{Q}_{[i']})$ is the KL information for the nth sample $\boldsymbol{x}_n$, where $\boldsymbol{P}_{[i']} = [P_n^1, \cdots, P_n^{i'-1}, P_n^{i'+1}, \cdots, P_n^K]^{\mathrm{T}}$ and $\boldsymbol{Q}_{[i']} = [Q_n^1, \cdots, Q_n^{i'-1}, Q_n^{i'+1}, \cdots, Q_n^K]^{\mathrm{T}}$, and is obtained using the probability variable $\boldsymbol{x}_{[i']}$ in which the i'th class is eliminated. In addition, N is the total number of samples, and N' is total after the class elimination. If the partial KL information $E_{[i']}$ is greater than 1, the i'th class is difficult to classify so needs to be eliminated.

B. Class Selection Method

In this paper, the pdf of sample data set $\boldsymbol{x}^*$ is estimated through PNN learning based on minimization of the KL information. In addition, using the data set $\boldsymbol{x}_{[i']}^*$ for which the i'th class is eliminated from $\boldsymbol{x}^*$, the partial KL information measure can be obtained at the same time as the learning network. Class selection is performed as follows by introducing the cross-validation method [12]:

(I) The vector set $\boldsymbol{x}^* \in \Re^L$ $(|\boldsymbol{x}^*| = N)$ is separated into M subsets of approximately equal size; the mth subset is $\boldsymbol{y}_m^*$, and the others are $\overline{\boldsymbol{y}}_m^*$ $(m = 1, 2, \cdots, M,$ $|\boldsymbol{y}_m^*| = S_m,$ $|\overline{\boldsymbol{y}}_m^*| = N - S_m,$ $\boldsymbol{y}_m^* \cup \overline{\boldsymbol{y}}_m^* = \boldsymbol{x}^*)$. The subset index m is set to one $(m = 1)$ and the
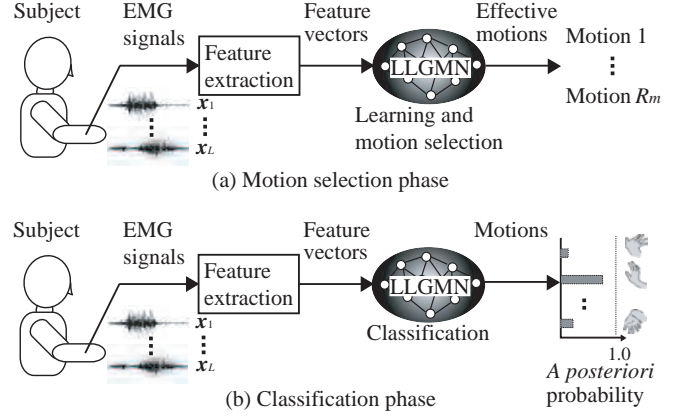

(a) Motion selection phase

(b) Classification phase

Fig. 1. Structure of the proposed motion selection method

subsequent processes are repeated under the M-fold cross-validation method [12].

(II) It is defined that $r_m = 0$, $R_m = K$, and that $\boldsymbol{y}_{[i']m}^*$ and $\overline{\boldsymbol{y}}_{[i']m}^*$ form the vector set for which the i'th class is eliminated from the vector set formed by $\boldsymbol{y}_m^*$ and $\overline{\boldsymbol{y}}_m^*$, respectively.

(III) The KL information $I(\boldsymbol{P}, \boldsymbol{Q}|\overline{\boldsymbol{y}}_m^*)$ and $I(\boldsymbol{P}_{[i']}, \boldsymbol{Q}_{[i']}|$ $\overline{\boldsymbol{y}}_{[i']m}^*)$ are computed using $\overline{\boldsymbol{y}}_m^*$ and $\overline{\boldsymbol{y}}_{[i']m}^*$. From this calculation, the partial KL information $E_{[i']m}$ for the data sets $\overline{\boldsymbol{y}}_m^*$ and $\overline{\boldsymbol{y}}_{[i']m}^*$ (Equation (2)) is obtained $(i' = 1, 2, \cdots, R_m)$. Moreover, $\boldsymbol{x}_{[i']}^* = \overline{\boldsymbol{y}}_{[i']m}^* \cup \boldsymbol{y}_{[i']m}^*$ is used to ascertain the estimated accuracy of the model $F_{[i']m}$ which is the average classification rate for $\boldsymbol{x}_{[i']}^*$.

(IV) Using the obtained $E_{[i']m}$, the jth class for which the partial KL information $E_{[j]m}$ is the largest, in other words, the most difficult one to classify, is obtained.

(V) If $F_{[j]m} < F_\theta$ (where F_θ is a threshold for the termination condition), and the jth class is eliminated. Then, let $\overline{\boldsymbol{y}}_m^* = \overline{\boldsymbol{y}}_{[j]m}^*$, $\boldsymbol{y}_m^* = \boldsymbol{y}_{[j]m}^*$, $r_m = r_m + 1$, and $R_m = K - r_m$ (where $r_m \leq K - 1$), and the process returns to Step (III). Otherwize ($F_{[j]m} \geq F_\theta$ or $r_m = K$), the process goes to Step (VI).

(VI) If $m < M$, it is set that $m = m + 1$, and the process is repeated from Step (II). Meanwhile, if $m = M$, the selection is terminated, and the model (r_m is the minimum) for which $F_{[j]m}$ is the maximum and $I(\boldsymbol{P}_{[j]}, \boldsymbol{Q}_{[j]}|\overline{\boldsymbol{y}}_{[j]m}^*)$ is the minimum is the best.

Eliminating classes that are difficult to classify one by one through this procedure enables the selection of effective classes (i.e., R_m classes).

III. MOTION SELECTION FOR EMG CLASSIFICATION

To accurately control machines such as EMG-based prostheses using EMGs, accurate classification of user motions from EMGs is necessary. As the number of motions for classification increases, however, classification problems become more complex so that the accuracy of classification

may decrease. Thus, a novel method of selecting appropriate motions (meaning that users stably generate EMGs) is necessary. This section reports on the use of a partial KL information-based class selection method in its application to motion selection for EMG classification. The structure of the motion selection method is shown in Fig. 1. In the proposed method, the subject's motions are learned by the LLGMN [5], and effective motions for classification are selected at the same time. The subject's motions can therefore be accurately classified using the selected motion only.

A. Feature Extraction

First, EMG signals measured from L pairs of electrodes are digitized using an A/D converter (sampling frequency: 1 [kHz]), and are rectified and filtered out through a second-order Butterworth filter (cut-off frequency: 1 [Hz]). These sampled signals are defined as $EMG_l(n)$ $(l = 1, \cdots, L)$. Next, the $EMG_l(n)$ values are normalized to make the sum of L channels equal to 1:

$$x_l(n) = \frac{EMG_l(n) - \overline{EMG_l^{st}}}{\sum_{l'=1}^{L} \left(EMG_{l'}(n) - \overline{EMG_{l'}^{st}} \right)}, \quad (3)$$

where $\overline{EMG_l^{st}}$ is the mean value of $EMG_l(n)$ while the muscles are relaxed.

B. Learning and Motion Selection

The feature vector $\boldsymbol{x}(n)$ $(n = 1, \cdots, N)$ is first input to the LLGMN, which is based on the GMM and a log-linear model of the pdf [5]. By applying the log-linear model to a product of a mixture coefficient and the mixture component of the GMM, the semiparametric pdf model is incorporated into a three-layer feedforward NN. This network can calculate the *a posteriori* probability $\boldsymbol{Y}(n) = [Y_1(n), Y_2(n), \cdots, Y_K(n)]^{\mathrm{T}}$ for the input pattern $\boldsymbol{x}(n)$ after learning [5].

A simple backpropagation algorithm is feasible when the teacher vector $\boldsymbol{T}(n) = [T_1(n), \cdots, T_k(n), \cdots, T_K(n)]^{\mathrm{T}}$ for the nth input vector $\boldsymbol{x}(n)$ is given. As the energy function J for the network,

$$J(\boldsymbol{x}^*) = I(\boldsymbol{T}, \boldsymbol{Y}|\boldsymbol{x}^*) = \frac{1}{N} \sum_{n=1}^{N} \sum_{k=1}^{K} T_k(n) \log \frac{T_k(n)}{Y_k(n)} \quad (4)$$

is used, and learning is performed to minimize the KL information for true distribution. This means that a learned network can approximate the data distribution. Using this energy function, therefore, it is possible to select classes (motions) that can be classified at the same time as the learning network.

The initial vector set $\boldsymbol{x}^*$ consists of feature vector $\boldsymbol{x}(n)$ $(n = 1, \cdots, N)$ calculated from the EMG signals measured, and classes (motions) that are difficult to classify are eliminated one by one from $\boldsymbol{x}^*$ based on the class selection method described in Section II-B. Here, the estimation accuracy of the pdfs in the model $F_{[i']m}$ is defined as the average classification rate for which $\boldsymbol{x}_{[i']}^* = \boldsymbol{y}_{[i']m}^* \cup \overline{\boldsymbol{y}}_{[i']m}^*$ is classified (the classification method is described in Section
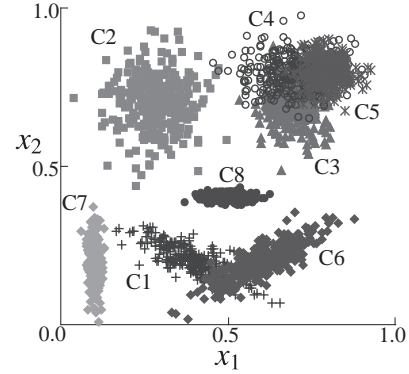
Fig. 2. Examples of artificial data for experiments

class	μ_{x_1}	μ_{x_2}	σ_{x_1}	σ_{x_2}	$\sigma_{x_1 x_2}$
C1	0.4	0.2	0.08	0.05	-0.8
C2	0.3	0.7	0.07	0.08	0
C3	0.7	0.7	0.05	0.06	0
C4	0.7	0.8	0.08	0.05	0
C5	0.8	0.8	0.04	0.04	0
C6	0.6	0.2	0.08	0.05	0.8
C7	0.1	0.2	0.01	0.06	-0.1
C8	0.5	0.4	0.04	0.01	0

III-C). After motion selection, r_m is the number of classes eliminated and R_m is the number of classes obtained.

C. Motion Classification

For EMG classification, since the output of the network's third layer provides the *a posteriori* probability of each class, Bayesian discrimination (in which the class with the highest *a posteriori* probability becomes the result) is used for classification. Here, for accurate discrimination, entropy is defined with the following equation:

$$H(n) = -\sum_{k=1}^{R_m} Y_k(n) \log Y_k(n). \quad (5)$$

Entropy indicates the ambiguity of each *a posteriori* probability, and if it is greater than the pre-specified discrimination threshold H_d, the discrimination is suspended.

IV. EXPERIMENTS

A. Simulation Experiments

1) Experimental Conditions: Fig. 2 shows a scatter diagram of the artificial data used in the experiments. The data consist of eight classes ($K = 8$) in a two-dimensional input space, and each class consists of one normal distribution with parameters as shown in Table I. In the experiments, 120 data for each class were used as an initial data set $\boldsymbol{x}^*$ for the class selection method (see Section II-B), and six-fold cross-validation was carried out ($M = 6$). This means that the

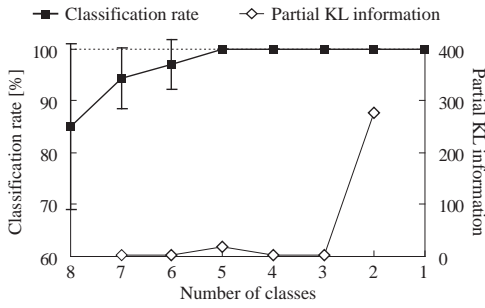Fig. 3. Resulted classification rates and partial KL information in the simulation experiments

Fig. 5. Overview of the sheet electrode

TABLE II

SELECTED CLASSES AND CLASSIFICATION RATES [%] IN THE
SIMULATION EXPERIMENTS

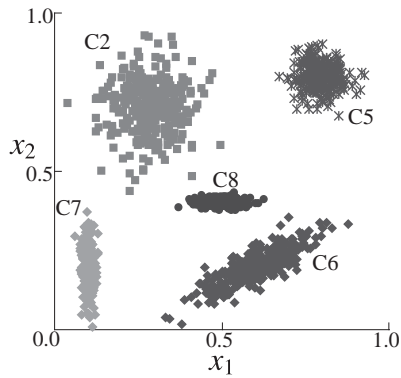| Selected classes		2, 5, 6, 7, 8
Classification rate of learning and test data	All classes	85.0
	Selected classes	100
Classification rate of evaluation data	All classes	86.9
	Selected classes	99.8

Fig. 6. Forearm motions for classification

Fig. 4. Scatter diagram of selected classes in the simulation experiments

number of learning data $|\overline{\boldsymbol{y}}_m^*|$ was 100 and the number of test data $|\boldsymbol{y}_m^*|$ was 20 for each class. In addition, the termination condition F_θ was set to 100 %. Moreover, 180 data for each class were used as evaluation data.

2) Results and Discussion: Fig. 3 shows the relationship between classification rates and partial KL information when classes are eliminated one by one using the proposed method. Note that the dashed line represents the threshold value of the termination condition ($F_\theta = 100$ %). It is confirmed that classification rates are lower than the threshold with eight to six classes, and five classes were selected using the termination condition of the proposed method. Table II lists the selected classes and their classification rates, respectively,
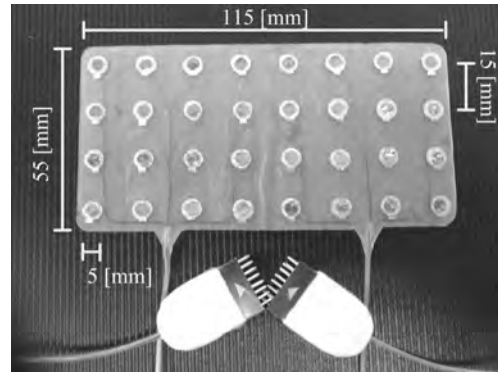
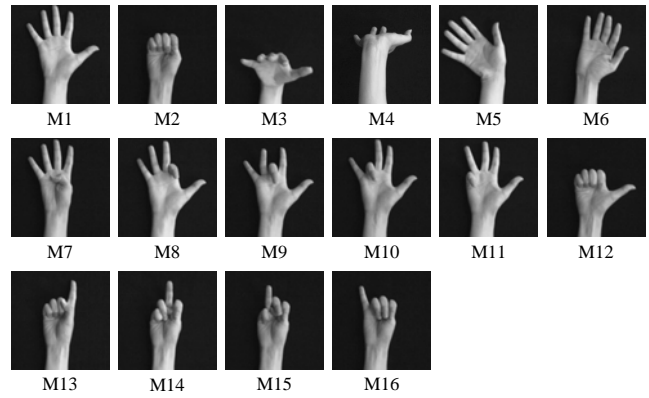and Fig. 4 shows a scatter diagram of the selected classes. From the results, it is confirmed that the overlap classes (C1, C3 and C4) are eliminated with increasing the classification rates. The classification rates for the selected classes are 12.9 % higher than that for all classes, and the average classification rate is 99.8 ± 0.4 % examined by the evaluation data. The results indicated that the proposed method was capable of accurate classification through one by one elimination of classes that are difficult to classify.

B. Motion Selection Experiments

To verify the effectiveness of the proposed class selection method, we performed EMG-based motion classification experiments.

1) Experimental Conditions: The subjects were three healthy volunteers (A – C). In the experiments, a pair of sheet electrodes (see Fig. 5) were attached to their right forearms ($L = 13$). Sheet electrode (Unique Medical Co., Ltd. [14]) is a type of electrode array referencing a subdural electrode. The center-to-center distance between adjacent electrodes is 15 [mm], where each electrode is 4 [mm] in diameter, and each sheet has $4 \times 8 = 32$ platinum electrodes. The subjects performed each of 16 motions ($K = 16$; hand grasping (M1), hand opening (M2), wrist flexion (M3) and extension
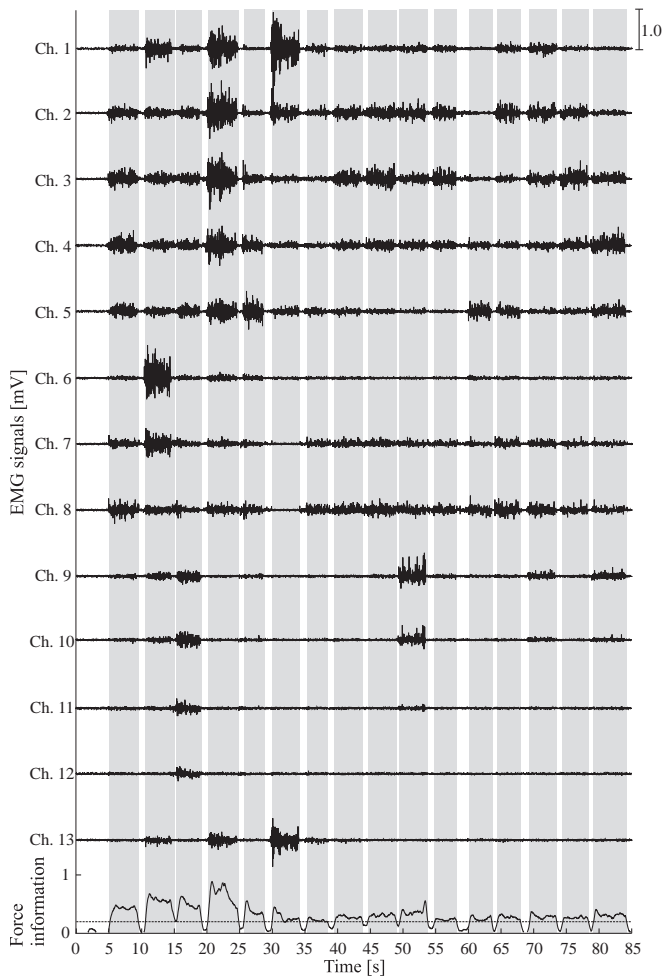
Fig. 7. Examples of measured signals in the motion selection experiment (Subject C)

Fig. 8. Relationships between classification rates and partial KL information in the motion selection experiment (Subject A)

(M4), ulnar flexion (M5), radial flexion (M6), each finger flexion (M7 – M11) and extension (M12 – M16); see Fig. 6) for a few seconds. The number of trial was 15, and data were measured using an MT-11 multitelemeter (gain: 16 [db]; NEC Sanei Co., Ltd.). Since the maximum number of channels which can measured using MT-11 is 13, 13 pairs of electrodes were selected in advance from all the electrodes in the sheets.

The sampling frequency was 1 [kHz], six sets of measured data were used for learning and motion selection (referred to here as the motion selection data), and the others were used only for classification (referred to as evaluation data). The number of motion selection data was 120 samples (20 randomly selected samples from 1,000 samples for each motion measured in each trial), and the number of evaluation data for each motion was 9,000 samples. The motion selection data were also divided into six subsets ($M = 6$). These data were taken as the initial vector sets, and motions were evaluated.

2) Results and Discussion: To evaluate whether appropriate motions for classification can be selected, a motion classification and selection experiment was performed. Fig.
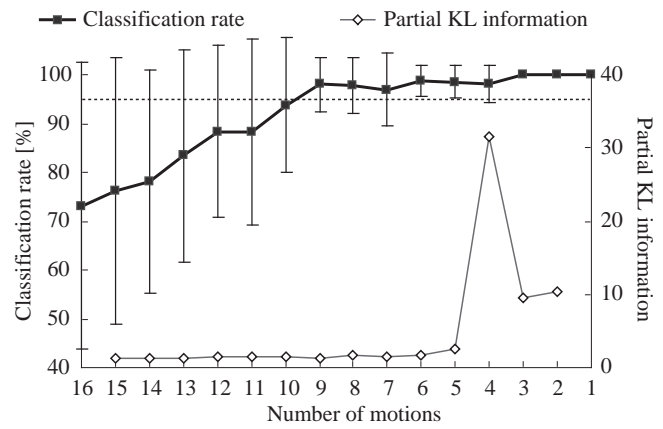
7 shows an example of the measured data, and force information. The gray zone indicate that the force information greater than the predefined threshold $F_{th} = 0.2$ [6], [7]. It can be seen that the EMG signals measured from each channel are different for each motion. Firstly, making the classification problem difficult, channels are selected using the channel selection method [10]. The results are shown in Table III, and the classification rates are low, especially for Subject A. In addition, Fig. 8 shows the relationship between the average classification rates for all classes and partial KL information when the classes were eliminated one by one using the proposed method. The dashed line represents the threshold value of the termination condition ($F_\theta = 95$ %). The figure shows the results of reducing the number of classes using motion selection data measured from Subject A. It is confirmed that the classification rates are lower than the threshold with 16 to 10 motions, and nine motions were selected by the termination condition of the proposed method. The average classification rate was $97.9 \pm 5.5\%$. Table IV and Fig. 9 show motions selected and the average classification rates of motion selection data and evaluation data for all subjects. Here, since the average classification rate of the motion selection data is high (99.0 ± 3.5 %), motions are not eliminated for Subject C. From these results, it is confirmed that classification rates increase: the average rates are 96.6 ± 1.9 % for motion selection data and 92.5 ± 0.9 % for evaluation data. These results lead us to the conclusion that the proposed method can be used to select appropriate motions for individual users conditions.

V. CONCLUSION

This paper proposes a class selection method using a new metric called the partial KL information measure to select effective classes. In order to confirm its effectiveness, the method was applied to motion selection for EMG classification. In experiments on forearm motion classification, appropriate motions were selected for each subject; the

	Sub. A	Sub. B	Sub. C
Selected channels	2, 3, 5, 9, 10	2, 3, 5, 8, 10, 13	2, 3, 4, 8, 9, 10, 11
Classification rate [%]	73.1	93.5	99.6

TABLE IV

MOTIONS SELECTED USING THE CLASS SELECTION METHOD. NOTE

THAT NO MOTIONS WERE ELIMINATED FOR SUBJECT C.

	Sub. A	Sub. B	Sub. C
Number of selected motions	9	14	16
Selected motions	1, 3, 4, 5, 6, 12, 13, 14, 15,	1, 2, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 15, 16	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16

average classification rate using the selected motions was 92.5 ± 0.9 % with a high accuracy.

In future research, we plan to enhance the method in order to allow the selection of channels [10] and motions at the same time. We further plan to apply this channel and motion selection approach based on the proposed partial KL information measure to training systems for EMG-based human-machine interfaces, and to create interactive training systems that can execute channel/motion selection and training of EMG control ability of the user.

VI. ACKNOWLEDGMENTS

(a) Classification rates of selection data

(b) Classification rates of evaluation data

Fig. 9.   Classification rates using all motions and selected motions

REFERENCES

[1] D. Graupe, J. Salahi, and D. Zhang, "Stochastic Analysis of Myoelectric temporal Signatures for Multifunctional Single-site Activation of Prostheses and Orthoses", *Journal of Biomedical Engineering*, vol. 7, no. 1, 1985, pp. 18-29.

[2] J. Cao, D. B. Sanders, "Multivariate Discriminant Analysis of the Electromyographic Interference Pattern: Statistical Approach to Discrimination Among Controls, Myopathies and Neuropathies", *Medical and Biological Engineering and Computing*, vol. 34, no. 5, 1996, pp. 369-374.

[3] HP Huang, CY Chen, "Development of a Myoelectric Discrimination System for a Multi-Degree Prosthetic Hand", *Proceedings of the 1999 IEEE International Conference on Robotics and Automation*, 1999, pp. 2392-2397.

[4] A. Subasi, "EEG Signal Classification Using Wavelet Feature Extraction and a Mixture of Expert Model", *Expert Systems with Applications*, vol. 32, no. 4, 2007, pp. 1084-1093

[5] T. Tsuji, O. Fukuda, H. Ichinobe, and M. Kaneko, "A Log-Linearized Gaussian Mixture Network and Its Application to EEG Pattern Classification", *IEEE Transactions on Systems, Man, and Cybernetics-Part C*, vol. 29, no. 1, 1999, pp. 60-72.

[6] O. Fukuda, T. Tsuji, M. Kaneko, and A. Otsuka, "A Human-Assisting Manipulator Teleoperated by EMG Signals and Arm Motions", *IEEE Transactions on Robotics and Automation*, vol. 19, no. 2, 2003, pp. 210-222
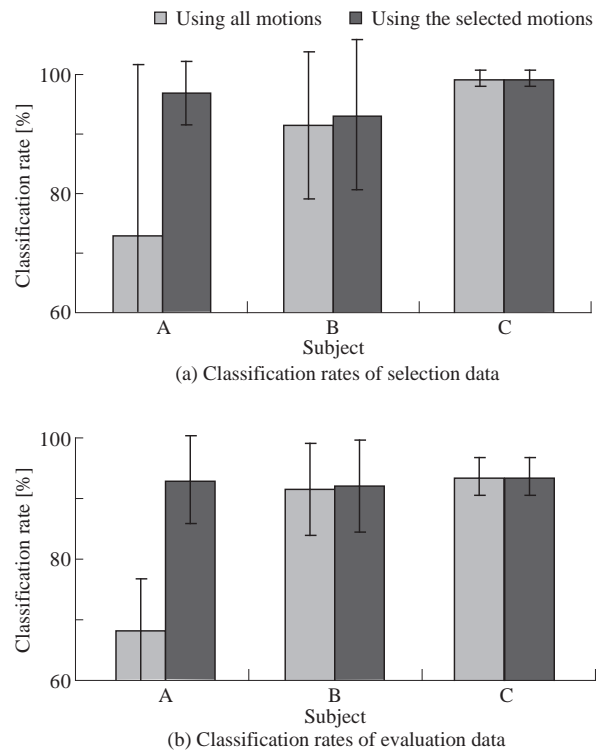
[7] H. Ogino, J. Arita, and T. Tsuji, "A Wearable Pointing Device using EMG Signals", *Journal of Robotics and Mechatronics*, vol. 17, no. 2, 205, pp. 173-180

[8] C. F. Olson, "Parallel Algorithms for Hierarchical Clustering", *Parallel Computing*, vol. 21, no. 8, 1995, pp. 1313-1325.

[9] J. Macqueen, "Some Methods for Classification and Analysis of Multivariate Observations", *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, 1967, pp 281-297.

[10] T. Shibanoki, K. Shima, T. Tsuji, A. Otsuka and T. Chin, "A Novel Channel Selection Method Based on Partial KL Information Measure for EMG-based Motion Classification", *ICBME2008 - 13th International Conference on Biomedical Engineering* (IFMBE Proceedings, Vol. 23), 2008, pp. 694-698.

[11] S. Kullback, R. A. Leibler, "On Information and Sufficiency, Annals of Mathematical Statistics", vol. 22, no. 1, 1951, pp. 79-86.

[12] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection", *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, vol. 2, no. 12, 1995, pp. 1137-1143.

[13] R. Nath, R. Pavur, "A New Statistic in the One-way Multivariate Analysis of Variance", Computational Statistics and Data Analysis, vol. 2, no. 4, 1985, 297-315.

[14] Unique Medical Co., Ltd.: http://www.mmjp.or.jp/unique-medical/