Proceedings of the 2012 ICME
International Conference on Complex Medical Engineering
July 1 - 4, Kobe, Japan

Multi-channel Surface EMG Classification Based on a Quasi-optimal Selection of Motions and Channels ¶

Taro SHIBANOKI*, Keisuke SHIMA†, Takeshi TAKAKI*, Yuichi KURITA*,
Akira OTSUKA‡, Takaaki CHIN§ and Toshio TSUJI*
* Graduate School of Engineering, Hiroshima University, Higashi-Hiroshima, Hiroshima, JAPAN
Email: {shibanoki, kurita, tsuji}@bsys.hiroshima-u.ac.jp, takaki@robotics.hiroshima-u.ac.jp
† Graduate School of Biomedical Science, Hiroshima University, Minami-ku, Hiroshima, Hiroshima, JAPAN
Email: shima@bsys.hiroshima-u.ac.jp
‡ Faculuty of Health and Welfare, Prefectural University of Hiroshima, Mihara, Hiroshima, JAPAN
Email: otuka@pu-hiroshima.ac.jp
§ Hyogo Rehabilitation Center, Nishi-ku, Kobe, Hyogo, JAPAN
Email: t-chin@pure.ne.jp

***Abstract*—This paper introduces a motion and channel selection method based on a partial Kullback-Leibler (KL) information measure. In the proposed method, the probability density functions of recorded data are estimated through learning involving a probabilistic neural network based on the KL information theory. Partial KL information is defined to support evaluation of the contribution of each dimension and class for classification. Effective dimensions and classes can then be selected by eliminating ineffective choices one by one based on this information, respectively. In the experiments, effective channels for classification were first selected for each of the six subjects, and the number of channels was reduced by 32.1 ± 25.5%. After channel selection, appropriate motions for classification were chosen, and the average classification rate for the motions selected using the proposed method was found to be 91.7 ± 2.5%. These outcomes indicate that the proposed method can be used to select effective channels and motions for accurate classification.**

***Index Terms*—Kullback-Leibler information, class selection, variable selection, pattern classification, electromyogram (EMG)**

I. Introduction

To control machines using biological signals, it is necessary to determine the relationships between these signals and actual motions. Various methods to classify biological signal patterns have been proposed [1]–[3]. Tsuji *et al.* proposed a log-linearized Gaussian mixture network (LLGMN) [2] that includes a Gaussian mixture model (GMM), and confirmed its effectiveness in biological signal classification [2], [3].

To classify biological signal patterns correctly, optimal recording sites for such signals and motions that are accurately classified must be selected in advance. Studies on optimal recording site selection have conducted by evaluating all combinations of EMG channels [4], [5]. In regard to selecting motions (classes) for classification, Kita *et al.* proposed a method for extracting motions that are clearly separated from one another [6]. With a large number of channels, however, computation takes a long time. Moreover, to realize motion classification using the selected channels' EMGs and/or motions, a classifier needs to be constructed after selection. Additionally, related efforts to date have focused only on the selection of optimal channels or motions.

The authors have developed a channel selection method based on the Kullback-Leibler (KL) information [7]. In this method, a partial KL information measure is defined as a metric of selecting optimal channels and ineffective channels for classification are eliminated one by one. A motion selection method involving the application of partial KL information for optimal class selection has also been proposed [8]. However, selecting the appropriate channels and motions cannot be performed at the same time.

Against such a background, this paper describes a motion and channel selection method based on the use of a partial KL information. In the proposed method, recorded signals are regarded as probability variables, and their probability density functions (pdfs) are estimated using probabilistic neural network (PNN) learning based on KL information. Additionally, ineffective channels and motions that are difficult to classify are eliminated one by one based on partial KL information [7], [8]. Under the proposed method, appropriate motions and channels can be selected at the same time.

II. A Motion and Channel Selection Method Based on a Partial KL Information Measure

A. Partial KL Information Measure [7], [8]

A pattern classification problem can be regarded as an estimation problem of the pdf for a given data [2]. 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

¶ This work was partially supported by a Grant-in-Aid for Scientific Research from the Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists (23·522) and the Japanese Ministry of Education, Science and Culture (22591634).

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

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

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.

Based on this concept, the authors proposed a partial KL information measure with $E_{[i]}$ and $E_{[i']}$ as a metric for variable and class selection [7], [8]:

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

$$E_{[i']} = \frac{I(\boldsymbol{P},\boldsymbol{Q})}{I(\boldsymbol{P'}_{[i']},\boldsymbol{Q'}_{[i']})} = \frac{N'}{N} \frac{\sum_{n=1}^{N} I_n(\boldsymbol{P},\boldsymbol{Q})}{\sum_{n'=1}^{N'} I'_n(\boldsymbol{P'}_{[i']},\boldsymbol{Q'}_{[i']})}, \quad (3)$$

Here, $I_n(\boldsymbol{P}_{[i]},\boldsymbol{Q}_{[i]})$ and $I_n(\boldsymbol{P}_{[i']},\boldsymbol{Q}_{[i']})$ are 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 $\boldsymbol{Q}_{[i']} = [Q_n^1, \cdots, Q_n^{i'-1}, Q_n^{i'+1}, \cdots, Q_n^K]^{\mathrm{T}}$. They are obtained using the probability variables $\boldsymbol{x}_{[i]}$ and $\boldsymbol{x}_{[i']}$, in which the ith dimension and the i'th class are eliminated, respectively. In addition, N is the total number of samples, and N' is the total after class elimination. The ith dimension or the i'th class for which $E_{[i]}$ or $E_{[i']}$ is the largest can be eliminated [7], [8].

B. Partial KL Information-based Motion and Channel Selection Using PNN

This section reports on motion and channel selection method using a partial KL information measure. The structure of the selection method is shown in Fig. 1. In the proposed method, the subject's motions are learned by the LLGMN [2], and effective channels and motions for classification are selected at the same time.

1) Feature Extraction [3], [7], [8]: 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, and resulted vector is defined as the feature vector $\boldsymbol{x}(n) \in \Re^L (n = 1, \cdots, N)$. Moreover, a channel average of $EMG_l(n)$ is defined as force information.

2) Learning and Selection of Channels and Motions [2], [3], [7], [8]: The LLGMN, which is based on the GMM and a log-linear model of the pdf [2], is used for learning. 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 [2].

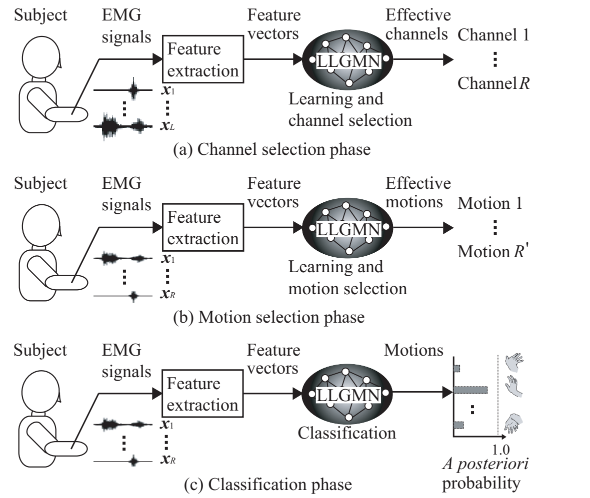

Fig. 1. Structure of the proposed motion and channel selection method

A simple backpropagation algorithm is feasible when the teacher vector $\boldsymbol{T}(n) = [T_1(n), T_2(n), \cdots, T_K(n)]^{\mathrm{T}}$ 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 optimal dimensions (channels) and classes (motions) that can be classified at the same time as the learning network.

The initial vector set $\boldsymbol{x}^* \in \Re^L$ is defined as $\boldsymbol{x}^* = \{\boldsymbol{x}(n) | n = 1, 2, \cdots, N\}$, and ineffective dimensions are eliminated one by one from $\boldsymbol{x}^*$ based on the variable selection method [7]. Here, D and $D_{[i]}$ are the average classification rates for $\boldsymbol{x}^*$ and $\boldsymbol{x}^*_{[i]}$, respectively, and $\boldsymbol{x}^*_{[i]}$ is the feature vector for which the ith dimension is eliminated from $\boldsymbol{x}^*$. Channels are first eliminated one by one while $D_{[i]} \geq D$. After channel selection, motions that are easy to classify can be selected using the class selection method [8] from the vector set $\boldsymbol{x}^{*'} \in \Re^R$, in which r dimensions are eliminated by the channel selection $(R + r = L)$. For motion selection, motions are eliminated one by one until the average classification rate for the selected motions $F_{[i']}$ is greater than the pre-specified threshold F_θ. Effective channels and motions are thus selected sequentially.

3) Motion Classification [2], [3], [7], [8]: 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 posteri-*

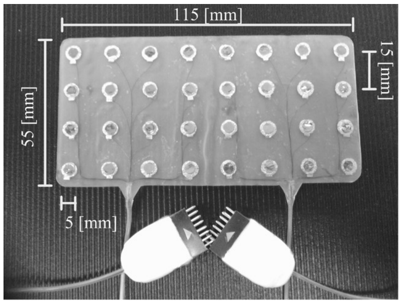

Fig. 2. Overview of the sheet electrode

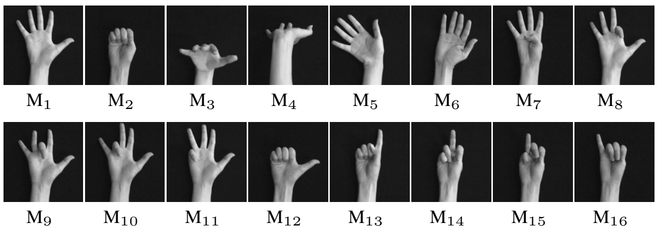

Fig. 3. Forearm motions for classification

ori probability becomes the result) is used for classification.

III. Experiments

A. Methods

The subjects were five healthy volunteers (A – E) and one upper-limb amputee (F). In the experiments, EMG signals were recorded using an MT-11 multitelemeter (gain: 60 [db]; NEC Sanei Co., Ltd.). A pair of sheet electrodes (Unique Medical Co., Ltd., see Fig. 2) was attached to each healthy subject's right forearm. 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 ($L = 13$). Two pairs of Ag/AgCl EMG electrodes (SEB120, GE Marquette) were attached to the amputation site of subject F ($L = 2$). Subjects A – E spent a few seconds performing each of 16 motions ($K = 16$; see Fig. 3) for 15 times, and Subject F performed 11 motions.

The sampling frequency was 1 [kHz], six sets of recorded data were used for learning and selection of channels and motions (referred to here as the selection data), and the others were used only for classification (referred to as evaluation data). The number of selection data was 120 samples (randomly selected from 1,000×6 samples for each motion), and the number of evaluation data for each motion was 9,000 samples.

B. Results and Discussion

Figure 4 shows an example of the recorded data, and force information. The gray zone indicate that the force information less than the predefined threshold F_{th}. The threshold value was set as $F_{\mathrm{th}} = 0.2$ to enable the detection of voluntary motions [3]. It can be seen that the EMG signals measured from each channel are different for each motion. Firstly, effective channels are selected for each user. Table I lists channel combinations chosen using the channel selection method, along with the average classification rates for selection data before and after channels were selected. It can be seen that the rates increased after channel selection for Subjects A–E. This is because redundant channels, which are not effective for classification, can be eliminated using this method. Additionally, since a small number of channels (two channels) were attached in advance, all two channels were selected for Subject F. It can be concluded that the channel selection method was capable for selecting effective channel combinations.

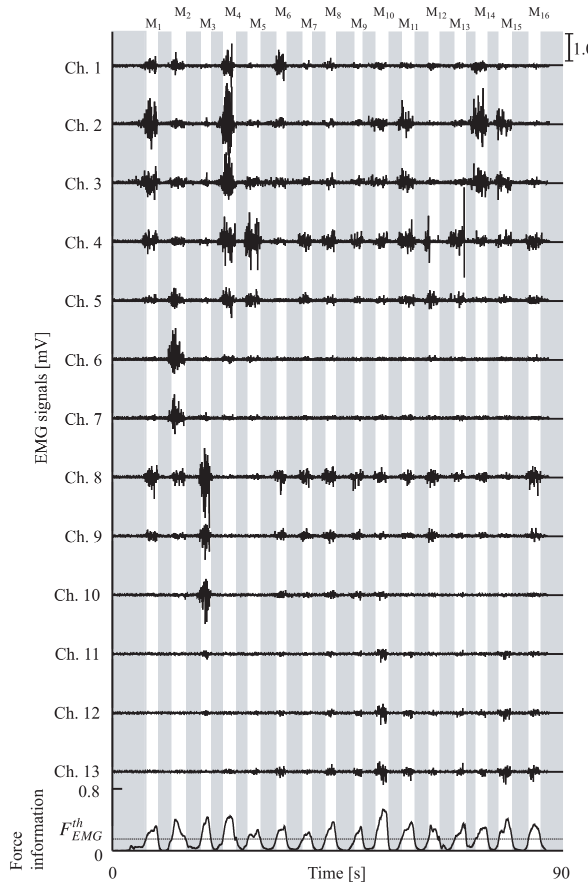

Fig. 4. Examples of recroded signals in the experiment

On the other hand, Fig. 5 shows the results of reducing the number of motions using selection data measured from Subject C. It is confirmed that the classification rates are lower than the threshold with 16 to 11 motions, and ten motions were selected by the termination condition of the proposed method. The average classification rate was 98.4 ± 2.1%. Table II shows motions selected and the average classification rates of selection data for all subjects. Here, since the average classification rate of the selection data is high (98.0 ± 1.2%), motions are not eliminated for Subject A, B and D. From these results, it is confirmed that classification

TABLE I
CHANNELS SELECTED USING THE CHANNEL SELECTION METHOD AND CLASSIFICATION RATES OF ALL MOTIONS FOR SELECTION DATA USING THE SELECTED CHANNELS. THIRTEEN CHANNELS WERE ATTACHED TO SUBJECTS A - E AND TWO CHANNELS TO SUBJECT F IN ADVANCE.

	Sub. A	**Sub. B**	**Sub. C**
Selected channels	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13	2, 3, 5, 8, 9, 10, 13	2, 3, 5, 9, 10
CR_{a} [%]	98.8	97.0	73.9
CR_{sc} [%]	99.4	97.4	79.4
	Sub. D	**Sub. E**	**Sub. F**
Selected channels	1, 2, 4, 5, 8, 11	2, 3, 4, 5, 6, 7, 9, 10, 11, 13	1, 2
CR_{a} [%]	95.2	92.6	43.9
CR_{sc} [%]	95.4	94.2	43.9

CR_{a}: Classification rates for the selection data using all channels
CR_{sc}: Classification rates for the selection data using the selected channels

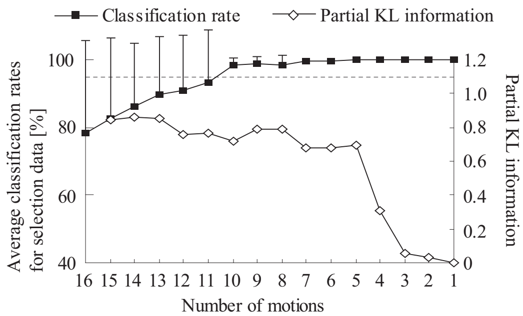

Fig. 5. Relationships between classification rates and partial KL information in the motion selection experiment (Subject C)

TABLE II
MOTIONS SELECTED USING THE PROPOSED MOTION SELECTION METHOD AND AVERAGE CLASSIFICATION RATES OF THE SELECTED MOTIONS FOR SELECTION DATA

	Sub. A	**Sub. B**	**Sub. C**
Selected motions	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	1, 3, 4, 5, 6, 9, 12 ,13 ,14, 15
CR_{sm} [%]	99.4	97.3	98.4
	Sub. D	**Sub. E**	**Sub. F**
Selected motions	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16	5, 6, 7, 10
CR_{sm} [%]	97.4	96.7	97.5

CR_{sm}: Classification rates for the selection data using the selected motions

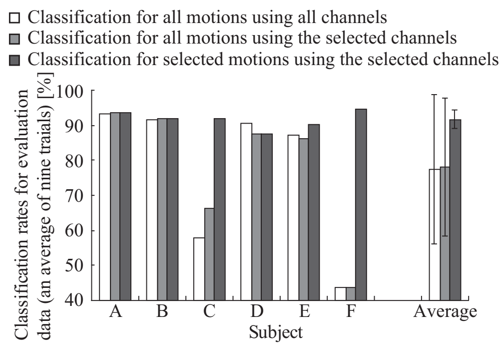

Fig. 6. Classification rates using all motions and selected motions. The average classification rates for all motions using all channels were shown as an comparison.

rates increase: the average rates are 97.8 $\pm$ 1.0% for selection data. Average classification rates for the evaluation data are also shown in Fig. 6, which outlines all subject's classification results obtained using the selected channels' EMGs for all motions and selected motions. The results of using all channels and motions are also shown for comparison. The average classification rate for all motions is $77.3 \pm 21.2\%$ when all channels are used, and $78.2 \pm 20.0\%$ when only the selected channels are used, and that for the selected motions is 91.7 $\pm$ 2.5%. The average rates do not differ when channels are eliminated and increase when motions are selected using the proposed method. These results indicate that the method can be used to select appropriate motions and channels for individual users' conditions.

IV. CONCLUSION

This paper introduces selection of appropriate motions and effective channels for classification using the partial KL information measure. In order to confirm its effectiveness, the method was applied to motion selection and EMG channel selection for EMG classification. In experiments on forearm motion classification, effective channels and appropriate motions were selected for each subject; the number of channels was reduced by $32.1 \pm 25.5\%$ and the average classification rate using the selected motions was 91.7 $\pm$ 2.5% with a high accuracy. In future research, we plan to apply this channel and motion selection approach to training systems for EMG-based human-machine interfaces.

ACKNOWLEDGMENT

We would like to cordially acknowledge and express our appreciation to Mr. Motonobu Shigeto and Mr. Kazuyuki Fujita for their assistance in the implementation of this research.

REFERENCES

[1] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," *Expert Systems with Applications*, vol. 32, no. 4, pp. 1084–1093, May 2007.
[2] T. Tsuji, O. Fukuda, H. Ichinobe, and M. Kaneko, "A log-linearized Gaussian mixture network and its application to EEG pattern classification," *IEEE Trans. on Systems, Man, and Cybernetics-Part C*, vol. 29, no. 1, pp. 60–72, 1999.
[3] O. Fukuda, T. Tsuji, M. Keneko, and A. Otsuka, "A human-assisting manipulator teleoperated by EMG signals and arm motions," *IEEE Trans. on Robotics and Automation*, vol. 19, no. 2, pp. 210–222, 2003.
[4] X. Zhang, X. Chen, Z. yan Zhao, Y. qiang Tu, J. hai Yang, V. Lantz, and K. qiao Wang, "Research on gesture definition and electrode placement in pattern recognition of hand gesture action SEMG," in *Medical Biometrics*, ser. Lecture Notes in Computer Science, 2007, pp. 33–40.
[5] A. Andrews, E. Morin, and L. McLean, "Optimal electrode configurations for finger movement classification using EMG," *Engineering in Medicine and Biology Society, 2009. 31st Annual International Conference of the IEEE*, pp. 2987–2990, 2009.
[6] K. Kita, R. Kato, and H. Yokoi, "Self-organized clustering approach for motion discrimination using EMG signal," *Proceedings of the 31st IEEE/EMBS Annual International Conference*, pp. 5044–5047, 2009.
[7] 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," *13th International Conference on Biomedical Engineering*, pp. 694–698, 2008.
[8] T. Shibanoki, K. Shima, T. Takaki, T. Tsuji, A. Otsuka, and T. Chin, "A class selection method based on a partial Kullback-Leibler information measure for biological signal classification," *System Integration (SII), 2010 IEEE/SICE International Symposium on*, pp. 317 –322, 2010.