Neuro-based Pattern Classification Method for EMG Signals not Belonging to Predefined Classes

Masaru Okamoto, Yukihiro Matsubara
Graduate School of Information Sciences,
Hiroshima City University
Hiroshima 731-3194
Japan
voice: [+81](82)830-1692
e-mail: {okamoto, matsubar}@hiroshima-cu.ac.jp
www: http://www.lake.its.hiroshima-cu.ac.jp/

Nan Bu
On-Site Sensing and Diagnosis Research Laboratory,
National Institute of Advanced Industrial Sciences and Technology
Tosu 841-0052
Japan
voice: [+81](942)81-4034; fax: [+81](942)81-3698
e-mail: bu@ieee.org
www: http://uist.aist.go.jp/kyushu/cie/

Toshio Tsuji
Graduate School of Engineering,
Hiroshima University
Higashi-Hiroshima 739-8527
Japan
voice: [+81](82)424-7677; fax: [+81](82)424-2387
e-mail: tsuji@bsys.hiroshima-u.ac.jp
www: http://www.bsys.hiroshima-u.ac.jp/

Abstract

This paper proposes a new Electromyograms (EMG) pattern classification method using prior probability of EMG signals. In this method, estimated prior probability based on Gaussian mixture model (GMM) is utilized for elimination of unexpected EMG signals. The method proposed automatically constructs GMM corresponding to complexity of EMG signals for training. Validity of the proposed method is shown by classification results of artificial data and EMG signals.

Keywords: pattern classification, EMG signals, prior probability, probabilistic neural network, unexpected data

1 Introduction

EMG pattern classification has been use to devise elaborate human-machine interfaces for people with physical disabilities [FTKO03]. In the recent years, various EMG pattern classification methods have been proposed [KYT06]. In particular, neural networks (NN) have been demonstrated as a promising classification tool, since their learning ability allows them to find optimum non-linear relationships between classes and feature patterns from data sets [KYT06], [HST89]. However, to effectively use NNs as the classifiers for applications, several problems, such as the choice of network structure, learning convergence and local minima, must be solved.

A probabilistic neural network (PNN), which estimates the probability density function of patterns, has been proven to be an efficient and important method for pattern classification. In particular, Tsuji et al. proposed a feed-forward PNN, a log-linearized Gaussian mixture network (LLGMN) based on the GMM and a log-linear model [TFIK99]. The LLGMN has been successfully applied to pattern classification of bioelectric signals [TSF+00], and has been used to develop human interface applications, such as prosthetic devices and EMG-based pointing devices [FTKO03], [TSF+00]. Moreover various pattern classification method based on LLGMN have been proposed. Bu and Tsuji have proposed a probabilistic neural network, a reduced-dimensional log-linearized Gaussian mixture network, for high-dimensional pattern classification [BT04]. Okamoto et al. have proposed several hierarchical pattern classification methods using LLGMN as classifier [OSMT08]. It is reported that these method have high classification performance for biological data such as EMG [BAT05] and hand shape signals [OSMT08].

However, LLGMN and the LLGMN-based classification methods suffer from inherent limitation of PNN, when dealing with EMG signals. In training procedure, it is assumed that all EMG signals belong to one of classes corresponding to outputs of PNN. For example, consider the case of three classes (grasping, wrist flexion and wrist extension) classification using EMG signals measured from forearm muscles. Although there are some EMG signals corresponding to other motion, such as hand opening, it is impossible for PNN to classify such signals into any class except three classes. To deal with this problem, one class classification method using Support Vector Machine (SVM) for elimination of outlier data have been proposed. In this method, Radial Basis Function is utilized as kernel of SVM [TD99]. Although this method removes outlier data different from given training data, there is no research report which has proposed multi-class classification using this approach and it takes a long time to decide fixed parameters of model through a trial and error process.

In this paper to deal with this problem, a novel EMG pattern classification method using prior distribution of EMG signals and probabilistic neural network (LLGMN) is proposed. By using the prior distribution based on GMM, the proposed method can remove unnecessary EMG signals (ex. other motions, outlier data). Moreover, the structure of prior distribution for EMG signals can be automatically estimated through a training procedure. After elimination, LLGMN can classify EMG signals into predefined classes. This procedure enables the proposed method to avoid the classification of unexpected EMG signals.

The rest of this paper is organized as follows. Section 2 proposed the details of the method of elimination of unexpected data and learning algorithm of the proposed structure. In Section 3, the EMG pattern classification method using LLGMN is provided. The results of computer simulation and phoneme pattern classification experiments of EMG signals are presented in Section 4. Finally, the last section concludes this paper.

2 Elimination of Unexpected Data

In the proposed method, the GMM are used in order to remove unnecessary data not belonging to predefined classes. The structure of GMM consists of some Gaussian distribution component. By estimating the number of components automatically, suitable structure for elimination can be constructed. The structure of the network and the constructing algorithm are explained below.

2.1 Structure

The prior probability $F(\boldsymbol{x})$ is given as

$$F(\boldsymbol{x}) = \sum_{m=1}^{M} \alpha_m g(\boldsymbol{x}, \boldsymbol{\mu}_m, \delta_m^2 \boldsymbol{E}), \tag{1}$$

where $\boldsymbol{x}$ is inputted data and M denotes the number of components, α_m is the mixture coefficient for component m, and $g(\boldsymbol{x}, \boldsymbol{\mu}_m, \delta_m^2 \boldsymbol{E})$ is a Gaussian distribution with mean vector $\boldsymbol{\mu}_m$and covariance matrix $\delta_m^2 \boldsymbol{E}$. $\boldsymbol{E}$ is the identity matrix. If $F(\boldsymbol{x})$ is greater than the threshold Tp, the data is classified into predefined classes by LLGMN. Otherwise, the classification is suspended.

2.2 Learning Algorithm

The proposed method can automatically estimate the suitable number of components corresponding to the complexity of training data. In the training procedure, a set of vectors $(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N)$ are utilized. The details of the proposed training scheme is as follows:
Step 1 Initialization:

1. Set the number of components M as 1 and the termination threshold $\delta^{(p)}$ as any given real number.

2. Initialize the mean vector $\boldsymbol{\mu}_1$ with randomized values, and then set δ_1^2 as the maximum value of the ii element of Σ calculated from the following equation and α_1 as 1.

$$\Sigma = \frac{1}{N}\sum_{n=1}^{N}(\boldsymbol{x}_n - \boldsymbol{\mu}_1)(\boldsymbol{x}_n - \boldsymbol{\mu}_1)^T \qquad (2)$$

Step 2 Update the mean vector:

1. Set the training iteration t as 1.

2. Update the mean vectors according to the following equations for all training data [Koh95].

$$\Delta\mu_{m'}(\boldsymbol{x}_n) = (1 - \frac{t}{T})(\boldsymbol{x}_n - \boldsymbol{\mu}_{m'}) \qquad (3)$$
$$m' = \arg\max_m g(\boldsymbol{x}_n, \boldsymbol{\mu}_m, \delta_m^2 \boldsymbol{E}) \qquad (4)$$

   where T is the predefined maximum iteration number and $m = 1, \cdots, M$.

3. Classify training data into M groups decided from Equation (4).

4. Compute δ_m^2 and the mixture coefficient α_m according to Equations (5) and (6) [MN99].

$$\delta_m^2 = \max_i \delta_{ii}^{(m)} \qquad (5)$$
$$\alpha_m = \frac{||G_m||}{N} \qquad (6)$$

   where $\delta_{ij}^{(m)}$ is the ij element of Σ_m (see Equation (7)), G_m is the set of data clasified into group m and $||G_m||$ is the number of data belonging to group G_m.

$$\Sigma_m = \frac{\sum_{\boldsymbol{x}_n \in G_m}(\boldsymbol{x}_n - \boldsymbol{\mu}_m)(\boldsymbol{x}_n - \boldsymbol{\mu}_m)^T}{||G_m||} \qquad (7)$$

5. This step of training repeats, until t reaches a predefined number T.

Step 3 Addition of component:

1. Stop the training procedure, if $\delta_m^2 > \delta^{(p)} - \epsilon$ and $M > 1$, where ϵ is a small positive number.

2. Otherwise, a component is added. $\delta^{(p)}$ is set as the maxmum variance δ_m^2 for next validation. Then, the mean vectors of the added component is initialized randomly and δ_m^2 and α_m of all components are calculated according to Equations (5) and (6). Then, go to Step 2.

3 EMG Pattern Classification Method

The proposed EMG pattern classification consists of three parts: (1) EMG feature extraction, (2) elimination of unexpected EMG signals and (3) classification network (LLGMN).

L channels of EMG signals are recorded using surface electrodes attached to muscles. The EMG signals are measured with a sampling frequency $f = 1000$Hz, then rectified and filtered by a Butterworth filter (cutoff frequency: 1Hz). Each sampled EMG pattern, defined as $\boldsymbol{EMG}(t)$ was normalized to make the sum of L channels equal to 1 using the following equation,

$$x_l^{(t)} = \frac{EMG_l(t) - EMG_l^{st}}{\sum_{l'=1}^{L}(EMG_{l'}(t) - EMG_{l'}^{st})}, \qquad (8)$$

where EMG_l^{st} is the mean value of $EMG_l(t)$ measured while relaxing the muscles. The feature vectors $\boldsymbol{x}(t) = [x_1(t), x_2(t), \cdots, x_L(t)]$ are inputted into classification network. A power level is estimated from the EMG signals as

$$power(t) = \frac{1}{L}\sum_{l=1}^{L}\frac{EMG_l(t) - EMG_l^{st}}{EMG_l^{max} - EMG_l^{st}}, \qquad (9)$$

where EMG_l^{max} is the mean value of $EMG_l(t)$ measured under the maximum voluntary contraction. The power level is compared with a prefixed threshold M_d to determine whether the motion actually happened.

For elimination of unexpected EMG signals, a prior probability calculated from GMM described

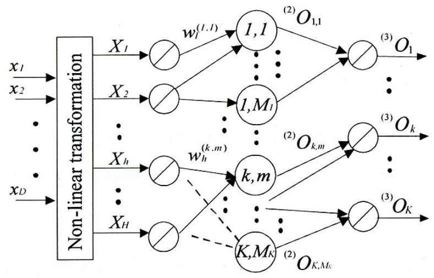


Figure 1: Structure of LLGMN

in Section 2 is employed. EMG signals corresponding to predefined classes are classified by LLGMN. The structure of LLGMN is shown in Fig. 1. In order to represent a normalized distribution corresponding to component of GMM as weight coefficient, the input EMG signal $\boldsymbol{x}(t)$ is converted into the modified input vector $\boldsymbol{X}$ as follows:

$$\boldsymbol{X} = \{1, \boldsymbol{x}^T, x_1^2, x_1x_2, \cdots, x_2^2, \cdots, x_2x_L, \cdots, x_L^2\}^T \tag{10}$$

The first layer of LLGMN consists $H = 1 + L(L+3)/2$ units, which correspond to the dimension of the input vector $\boldsymbol{X}$, and the identity function is used for the activation function of each unit. The outputs of the first layer multiplied by weight $w_h^{(k,m)}$ are transmitted to the second layer. Where $w_h^{(K,M_K)} = 0$, K and M_K denote the number of classes and components belonging to class M, respectively. In this layer, LLGMN calculates the posteriori probability of each Gaussian component $\{k, m\}$. The unit k in the third layer integrates the outputs of M_k units in the second layer.

$$^{(2)}I_{k,m}(t) = \sum_{k=1}^{H} {}^{(1)}O_h(t) w_h^{(k,m)} \tag{11}$$

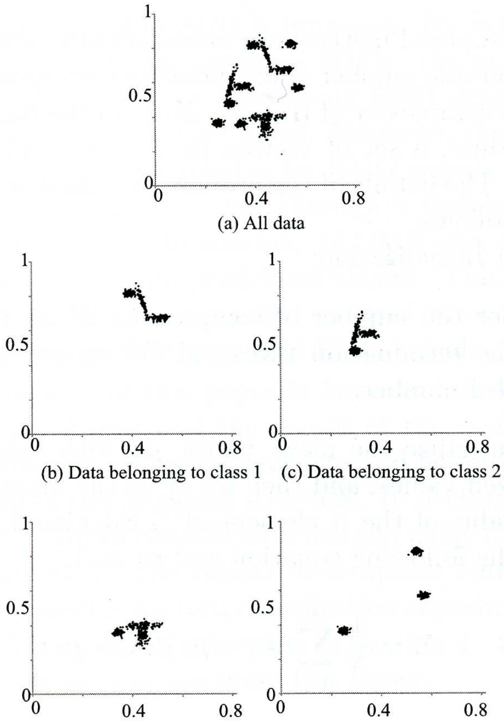
$$^{(2)}O_{k,m}(t) = \frac{\exp(^{(2)}I_{k,m}(t))}{\sum_{k'=1}^{K}\sum_{m'=1}^{M_{k'}} \exp(^{(2)}I_{k',m'}(t))} \tag{12}$$

The relationship between the input $^{(3)}I_k(t)$ and the output $^{(3)}O_k(t)$ in the third layer is

$$^{(3)}O_k(t) = {}^{(3)}I_k(t) = \sum_{m=1}^{M_k} {}^{(2)}O_{k,m}(t) \tag{13}$$

The output of the third layer corresponds to the posterior probability $P(k|\boldsymbol{x}(t))$ of class k given the input vector $\boldsymbol{x}(t)$. The entropy of outputs is also calculated to prevent risk of misclassification. The entropy is defined as Equation 14.

$$H(\boldsymbol{x}(t)) = -\sum_{k=1}^{K} {}^{(3)}O_k(t) \log^{(3)} O_k(t) \tag{14}$$



(a) All data

(b) Data belonging to class 1 (c) Data belonging to class 2

(d) Data belonging to class 3 (e) Data not belonging to predefined classes.

Figure 2: Artificial data used for classification experiments

If the entropy H(x(t)) is less than the classification threshold Te, the specific motion with the largest probability is determined according to the Bayes' decision rule (shown in Equation 15). Otherwise, the determination is suspended.

$$Y(\boldsymbol{x}(t)) = \arg\max_k {}^{(3)}O_k(t) \tag{15}$$

4 Experiments

4.1 Numerical Experiments

First, pattern classification experiments on artificial data were conducted for evaluating the performance of the proposed method. A two-dimensional input space consisted of three classes ($K = 3$) and predefined class, each class and predefined class consisting of three components. Examples of the data are shown in Fig. 2. For each class, 200 samples were generated to train, and then the trained network was validated using test data (500 samples/class). The values of the parameters were $Tp = 0.01, \epsilon = 0.01$ and $te = 0.8$.

Fig. 3 shows the classification results by the proposed method for 10 independent trials (the

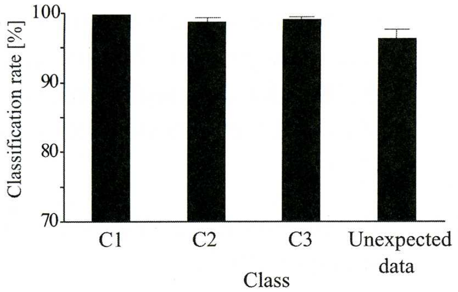


Figure 3: Classification results

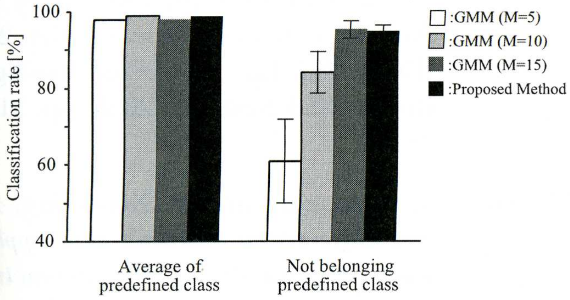


Figure 4: Comparison of elimination results

initial mean vectors and weights of LLGMN were chosen at random). The results indicate that the proposed method achieved the elimination of unexpected data and high classification performance.

In order to confirm the effectiveness of automatic addition of components, GMMs that fixed the number of components were used for comparison. In this experiments, test data for validation were divided into the data belonging to predefined classes and the unexpected data by the proposed method and the traditional GMMs (the number of components are 5, 10, 15). Fig. 4 shows the elimination results by the proposed method and the other methods using fixed GMMs. The mean value of the number of components for GMM was 17.6 ± 3.0. These results indicate that the proposed method can estimate successfully the number of components even for unexpected data not belonging to the predefined classes.

4.2 EMG pattern classifications

Phoneme classification based on EMG signals was conducted to examine performance of the proposed method. In the experiments, EMG signals measured from mimetic and cervical muscle were used to classify six Japanese phonemes

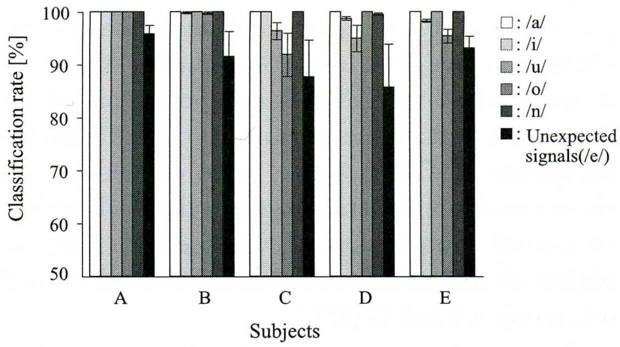


Figure 5: EMG classification results for five subjects

(/a/, /i/, /u/, /e/, /o/, /n/). In this experiment, classes corresponding to five phonemes (/a/, /i/, /u/, /o/, /n/) were used as predefined classes and EMG signals belonging to utterance /e/ were set as unexpected EMG signals. Five subjects (A, B, C, D and E) participated in the experiments.

Five pairs of Ag/AgCl electrodes (NT-511G: NIHON KOHDEN Corp.) were attached to the subject' s face (Depressor Anguli Oris, Zygomaticus Major, Masseter, Digastric, Depressor Labii Inferioris; a pair of electrodes on each muscle) with conductive paste. The EMG signals from five muscles were recorded (sampling frequency: 1kHz). The values of the parameters were $Tp = 0.01, \epsilon = 0.01, te = 0.8$ and $M_d = 0.25$.

Five sets of randomly chosen initial mean vectors and weights were used to train each sample data. The mean values and standard deviations of the classification rates are shown in Fig. 5. From this figure, it can be seen that the elimination of unexpected EMG signals and the classification accuracy of EMG signals belonging to predefined classes are achieved by using the prior probability of EMG signals and LLGMN for classification.

5 Conclusions

This paper proposes a new classification method for EMG signals using a prior probability of EMG signals estimated from GMM for the elimination of the EMG signals not belonging to predefined classes. In this method, the structure of GMM is automatically constructed by adding components to estimate the distribution of EMG signals for training. Moreover, by using LLGMN as classifier, classification part has high performance for adaptive learning based on statistical properties of EMG signals. To examine the classification capability and the accuracy of the proposed method, classification experiments of artificial data and phoneme pattern classification experiments of EMG signals were carried out. In these experiments, the proposed method achieved successful elimination of unexpected data and higher classification performance than other methods using a fixed GMM.

Our research group previously proposed the EMG-based Japanese speech synthesizer system [FFT05] based on LLGMN and a hidden Markov model. In this system, six Japanese phonemes are classified from EMG signals using a LLGMN. By using the method proposed in this paper as a classifier, it is expected that the speech synthesizer system can avoid some problems due to misclassification of unexpected EMG signals.

In future, we would like to improve the learning algorithm of GMM for elimination theoretically. It is also interesting to study the integration of elimination and classification.

References

[BAT05] N. Bu, J. Arita, and T. Tsuji, *A novel pattern classification method for multivariate emg signals using neural network*, Proc. of the First Int. Conf. on Natural Computation, Part II, 2005, pp. 165–174.

[BT04] N. Bu and T. Tsuji, *Multivariate pattern classification based on local discriminant component analysis*, Proc. of IEEE Int. Conf. on Robotics and Biomimetics, 2004, p. 290.

[FFT05] O. Fukuda, S. Fujita, and T. Tsuji, *A substitute vocalization system based on emg signals*, The IEICE Trans. on Information and Systems **J88-D-II** (2005), no. 1, 105–112 (in Japanese).

[FTKO03] O. Fukuda, T. Tsuji, M. Kaneko, and M. Otsuka, *A human-assisting manipulator teleoperated by emg signals and arm motion*, IEEE Trans. on Robotics and Automation **19** (2003), 210–222.

[HST89] A. Hiraiwa, K. Shimohara, and Y. Tokunaga, *Emg pattern analysis and classification by neural network*, IEEE Informational Conf. on System, Man and Cybernetics, 1989, pp. 1113–1115.

[Koh95] T. Kohonen, *Self-organizing maps*, Springer, 1995.

[KYT06] R. Kato, H. Yokoi, and A. Tamio, *Real-time learning method for adaptable motion-discrimination using surface emg signal*, Proc of the 2006 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, 2006, pp. 2127–2132.

[MN99] J. R. Magnus and H. Neudecker, *Matrix differential calculus with applications in statistics and econometrics*, Wiley, 1999.

[OSMT08] M. Okamoto, K. Shima, Y. Matsubara, and T. Tsuji, *Pattern discrimination method with boosting approach using hierarchical neural trees*, Proc. of the Institution of Mechanical Engineers, Part I, Journal of Systems and Control Engineering (in press) (2008).

[TD99] M. J. Tax and P. W. Duin, *Support vector domain description*, Pattern Recognition Letters **20** (1999), 1191–1199.

[TFIK99] 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: Applications and Reviews **29-1** (1999), 60–72.

[TSF+00] T. Tsuji, H. Shigeyoshi, O. Fukuda, H. Ichinobe, and M. Kaneko, *Biomimetic control of an extrenally powered prothetic forearm based on emg signals(in japanese)*, Trans. of the Japan Society of Mechanical Engineers **66** (2000), 2764–2771.