Journal of Intelligent Information Systems, 21:2, 113–126, 2003

EMG-Based Motion Discrimination Using a Novel Recurrent Neural Network

NAN BU bu@bsys.hiroshima-u.ac.jp
Department of Artificial Complex Systems Engineering, Hiroshima University, Higashi-Hiroshima, 739-8527, Japan

OSAMU FUKUDA fukuda.o@aist.go.jp
National Institute of Advanced Industrial Science and Technology, Tsukuba, 305-8564, Japan

TOSHIO TSUJI tsuji@bsys.hiroshima-u.ac.jp
Department of Artificial Complex Systems Engineering, Hiroshima University, Higashi-Hiroshima, 739-8527, Japan

Abstract. This paper presents a pattern discrimination method for electromyogram (EMG) signals for application in the field of prosthetic control. The method uses a novel recurrent neural network based on the hidden Markov model. This network includes recurrent connections, which enable modeling time series, such as EMG signals. Weight coefficients in the network can be learned using a well-known back-propagation through time algorithm. Pattern discrimination experiments were conducted to demonstrate the feasibility and performance of the proposed method. We were able to successfully discriminate forearm motions using the EMG signals, and achieved considerably high discrimination performance compared with other discrimination methods.

Keywords: neural networks, pattern discrimination, EMG, recurrent neural network

1. Introduction

A human-machine interface for prosthetic control is important in assisting the disabled who have lost manipulation capability of an upper limb. EMG signals are often used as an interface for prosthetic devices, since they are the manifestation of electrical stimulations, which motor units receive from the Central Nervous System (CNS), and indicate the activation level of motor units associated with muscle contractions. Different motions resulting from different modes of muscle activation generate different EMG patterns. Many researchers have tried to use EMG signals as a means for providing control commands for prosthetic arms (Graupe et al., 1978; Hiraiwa et al., 1989, 1992; Kelly et al., 1990; Tsuji et al., 1987, 1992, 1993, 1999; Huang and Chen, 1999; Fukuda et al., 2000).

Various techniques have been proposed to discriminate the EMG patterns. Graupe et al. (1978) used autoregressive (AR) models to represent the EMG signal, which was measured from a single electrode-site between the biceps and the triceps. Motion patterns were discriminated according to the parameters of the AR model. Tsuji et al. (1987) used a multi-dimensional AR model, and were able to discriminate forearm motions using frequency and amplitude characteristics extracted from multi-channel EMG signals. However, these methods did not achieve high discrimination performance, because they applied a linear model to approximate the nonlinear characteristics of the EMG signal, which varies significantly depending on factors such as muscle fatigue, sweat, and changes in electrode location.

The Neural Network (NN) is suitable for modeling nonlinear data, and is able to cover the distinction among different conditions. Several EMG pattern discrimination methods based on NN have been presented in the last decade. For example, Hiraiwa et al. (1989, 1992) used a back-propagation (BP) NN to perform pattern discrimination with frequency features. Kelly et al. (1990) succeeded in discriminating four forearm motions (flexion, extension, pronation, and supination) using a combination of BPNN and a Hopfield NN. Koike and Kawato (1994), Huang and Chen (1999), and others have also carried out similar works. However, BPNNs frequently used in the aforementioned studies do not offer high learning and discriminating performance, and a great deal of learning data, as well as a large number of learning iterations, is required.

Tsuji et al. (1992) proposed an NN including a statistical model, and used the entropy calculated from the outputs of the NN to improve discrimination (Tsuji et al., 1993). They also proposed a feedforward probabilistic NN, a Log-Linearized Gaussian Mixture Network (LLGMN) (Tsuji et al., 1999), which is based on the Gaussian mixture model (GMM) and the log-linear model of the probability density function. LLGMN was successfully applied to the EMG pattern classification, where eight motions of the forearm were classified using EMG signals measured by several pairs of electrodes (Fukuda et al., 2000). These methods, however, only focused on the static features of the EMG signal, and time-varying characteristics were not taken into account.

This paper proposes a new EMG pattern discrimination method using a recurrent NN, Recurrent Log-Linearized Gaussian Mixture Network (R-LLGMN) (Tsuji et al., 2001, 2003). R-LLGMN is based on the algorithm of the Hidden Markov Model (HMM) (Rabiner, 1989) and incorporates recurrent connections to make use of the time context in the EMG signal. With the weight coefficients well learned using a learning scheme of the back-propagation through time (BPTT) algorithm (Werbos, 1990), R-LLGMN can calculate the a posteriori probability of discriminating class for each EMG pattern. Pattern discrimination experiments of the EMG signal were conducted using R-LLGMN with five subjects, and the proposed method was compared with BPNN and LLGMN.

This paper is organized as follows. Section 2 explains the architecture and algorithm of the R-LLGMN. The EMG pattern discrimination method is described in Section 3. The discrimination experiments are presented in Section 4. The final section presents the conclusion of this paper.

2. Recurrent log-linearized Gaussian mixture network

The EMG signal is regarded stochastic in nature, which is the sum of the spike potential generated in the muscle fibers. To achieve an accurate estimation of the probability density function (pdf) of EMG signal, a semiparametric estimation method is adopted, because it has a flexible structure that can represent any distribution. In this method, the unknown distribution is defined as a weighted sum of a number of component distributions, where Gaussian distribution is often used. Moreover, to cope with the time-varying characteristic of EMG signal, we introduced the algorithm of HMM, which is a well-studied technology for temporal discrimination and has been successfully applied in the field of speech recognition. By applying the log-linear model, the component distribution in the probabilistic model is represented as the linear sum of the input vector and the coefficient vector so that we can incorporate them into the structure of NN, that is, the Recurrent Log-Linearized Gaussian Mixture Network (R-LLGMN) (Tsuji et al., 2001, 2003).

2.1. *HMM-based dynamic probabilistic model*

Let us consider a dynamic probabilistic model, as shown in figure 1, where there are C classes in this model and each class c ($c \in \{1, \ldots, C\}$) is composed of K_c states. Suppose that, for the given time series $\tilde{\mathbf{x}} = \mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(T)$ ($\mathbf{x}(t) \in \Re^d$), at any time $\mathbf{x}(t)$ must occur from one state k of class c in the model. With this model, the a posteriori probability for class c, $P(c \mid \tilde{\mathbf{x}})$, is calculated as

$$P(c \mid \tilde{\mathbf{x}}) = \sum_{k=1}^{K_c} P(c, k \mid \tilde{\mathbf{x}}) = \sum_{k=1}^{K_c} \frac{\alpha_k^c(T)}{\sum_{c'=1}^{C} \sum_{k'=1}^{K_{c'}} \alpha_{k'}^{c'}(T)}. \tag{1}$$

Here, $\alpha_t^c(k)$ is the forward variable, and it can be derived as

$$\alpha_k^c(1) = \pi_k^c b_k^c(\mathbf{x}(1)), \tag{2}$$

$$\alpha_k^c(t) = \sum_{k'=1}^{K_c} \alpha_{k'}^c(t-1)\gamma_{k',k}^c b_k^c(\mathbf{x}(t)) \quad (1 < t \leq T), \tag{3}$$

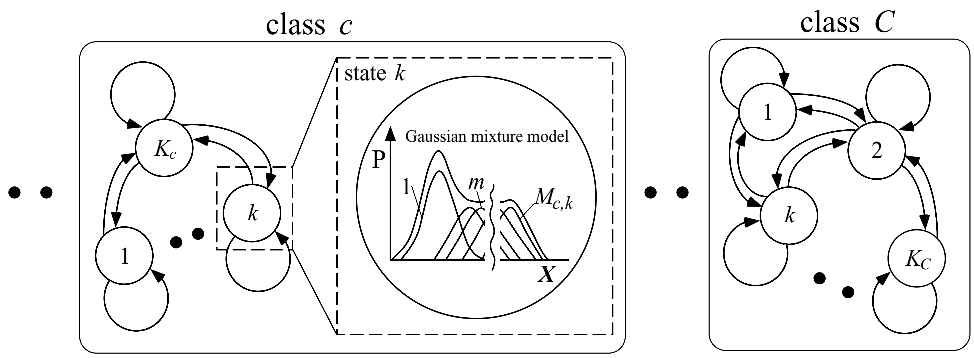

Figure 1. HMM-based dynamic probabilistic model with C classes and K_c states in class c.

where $\gamma^c_{k',k}$ is the probability of the state changing from k' to k in class c, and $b^c_k(\mathbf{x}(t))$ is defined as the a posteriori probability for state k in class c corresponding to $x(t)$. Also the a priori probability π^c_k is equal to $P(c,k)|_{t=0}$.

In this model, the a posteriori probability $b^c_k(\mathbf{x}(t))$ is approximated by summing up $M_{c,k}$ components of a Gaussian mixture distribution, and $\gamma^c_{k',k}b^c_k(\mathbf{x}(t))$ on the right side of (3) is derived in the form

$$\begin{aligned}\gamma^c_{k',k}b^c_k(\mathbf{x}(t)) &= \sum_{m=1}^{M_{c,k}} \gamma^c_{k',k} r_{c,k,m}\, g\big(\mathbf{x}(t); \boldsymbol{\mu}^{(c,k,m)}, \Sigma^{(c,k,m)}\big)\\ &= \sum_{m=1}^{M_{c,k}} \gamma^c_{k',k} r_{c,k,m} (2\pi)^{-\frac{d}{2}} \big|\Sigma^{(c,k,m)}\big|^{-\frac{1}{2}}\\ &\quad\times \exp\Bigg[-\frac{1}{2}\sum_{j=1}^{d}\sum_{l=1}^{j}(2-\delta_{jl})s_{jl}^{(c,k,m)}x_j(t)x_l(t)\\ &\quad+\sum_{j=1}^{d}\sum_{l=1}^{d}s_{jl}^{(c,k,m)}\mu_j^{(c,k,m)}x_l(t) - \frac{1}{2}\sum_{j=1}^{d}\sum_{l=1}^{d}s_{jl}^{(c,k,m)}\mu_j^{(c,k,m)}\mu_l^{(c,k,m)}\Bigg],\end{aligned} \tag{4}$$

where $r_{c,k,m}$, $\boldsymbol{\mu}^{(c,k,m)} = (\mu_1^{(c,k,m)}, \ldots, \mu_d^{(c,k,m)})^{\mathrm{T}}$, $\Sigma^{(c,k,m)} \in \Re^{d\times d}$, $s_{ij}^{(c,k,m)}$ and $x_i(t)$ stands for the mixing proportion, the mean vector, the covariance matrix of each component $\{c,k,m\}$, the element of the inverse of covariance matrix $\Sigma^{(c,k,m)-1}$, and the element of $\mathbf{x}(t)$, respectively.

The recurrent NN, R-LLGMN, is developed from the model defined above. For an input time series $\tilde{\mathbf{x}}$, the a posteriori probability for each class can be estimated with a well trained R-LLGMN. The network structure and the learning algorithm of R-LLGMN are explained in the following sections.

2.2. *Network structure*

R-LLGMN is a five-layer recurrent NN with feedback connections between the fourth and the third layers, the structure of which is shown in figure 2. First, the input vector series $\mathbf{x}(t) \in \Re^d$ $(t = 1, \ldots, T)$ is pre-processed into the modified input series $\mathbf{X}(t) \in \Re^H$ as follows:

$$\begin{aligned}\mathbf{X}(t) = \big(&1, \mathbf{x}(t)^{\mathrm{T}}, x, (t)^2, x, (t)x_2(t), \ldots, x, (t)x_d(t),\\ &x_2(t)^2, x_2(t)x_3(t), \ldots, x_2(t)x_d(t), \ldots, x_d(t)^2\big)^{\mathrm{T}},\end{aligned} \tag{5}$$

where the dimension H is determined as $H = 1 + d(d+3)/2$. The vector $\mathbf{X}(t)$ acts as the input of the first layer, and the identity function is used to activate each unit. The output of the hth $(h = 1, \ldots, H)$ unit in the first layer is defined as ${}^{(1)}O_h(t)$.

Unit $\{c, k, k', m\}(c = 1, \ldots, C; k', k = 1, \ldots, K_c; m = 1, \ldots, M_{c,k})$ in the second layer receives the output of the first layer, weighted by the coefficient $w^c_{k',k,m,h}$. The input

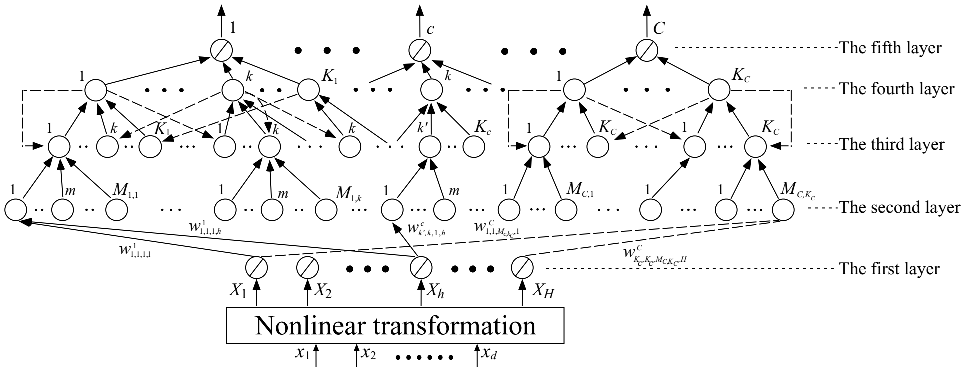

Figure 2. Structure of R-LLGMN.

${}^{(2)}I^c_{k',k,m}(t)$ and the output ${}^{(2)}O^c_{k',k,m}(t)$ are defined as

$$ {}^{(2)}I^c_{k',k,m}(t) = \sum_{h=1}^{H} {}^{(1)}O_h(t) w^c_{k',k,m,h}, \tag{6} $$

$$ {}^{(2)}O^c_{k',k,m}(t) = \exp\big({}^{(2)}I^c_{k',k,m}(t)\big), \tag{7} $$

where C is the number of discriminating classes, K_c is the number of states in class c, and $M_{c,k}$ denotes the number of GMM components in the state k of class c. In Eq. (7), the exponential function is used in order to calculate the probability of the input pattern.

The outputs of units $\{c, k, k', m\}$ in the second layer are added up and input into a unit $\{c, k, k'\}$ in the third layer. Also, the output of the fourth layer is fed back to the third layer. These are expressed as follows:

$$ {}^{(3)}I^c_{k',k}(t) = \sum_{m=1}^{M_{c,k}} {}^{(2)}O^c_{k',k,m}(t), \tag{8} $$

$$ {}^{(3)}O^c_{k',k}(t) = {}^{(4)}O^c_{k'}(t-1) {}^{(3)}I^c_{k',k}(t), \tag{9} $$

where ${}^{(4)}O^c_{k'}(0) = 1.0$ is for the initial phase. The recurrent connections between the fourth and the third layers play an important role in the process which corresponds to the forward computation (see Eq. (3)).

The activation function in the fourth layer is described as

$$ {}^{(4)}I^c_k(t) = \sum_{k'=1}^{K_c} {}^{(3)}O^c_{k',k}(t), \tag{10} $$

$$ {}^{(4)}O^c_k(t) = \frac{{}^{(4)}I^c_k(t)}{\sum_{c'=1}^{C} \sum_{k'=1}^{K_{c'}} {}^{(4)}I^{c'}_{k'}(t)}. \tag{11} $$

In the fifth layer, the unit c integrates the outputs of K_c units $\{c, k\}$ $(k = 1, \ldots, K_c)$ in the fourth layer. The relationship in the fifth layer is defined as:

$$^{(5)}I^c(t) = \sum_{k=1}^{K_c} {}^{(4)}O_k^c(t), \tag{12}$$

$$^{(5)}O^c(t) = {}^{(5)}I^c(t). \tag{13}$$

In the R-LLGMN, the a posteriori probability of each class is defined as the output of the last layer. After optimizing the weight coefficients $w_{k',k,m,h}^c$ between the first layer and the second layer, the NN can estimate the a posteriori probability of each class. Obviously, the structure of R-LLGMN is of good correspondence with the algorithm of HMM. R-LLGMN, however, is not just a copy of HMMs. The essential point of R-LLGMN is that the parameters in HMMs are replaced by the weight coefficients $w_{k',k,m,h}^c$, and this replacement removes restrictions of the statistical parameter in HMMs (e.g., $0\leq$ the transition probability ≤ 1, and standard deviations >0). Therefore, the learning algorithm of R-LLGMN is simplified and can be expected higher generalization ability than that of HMMs. That is the great advantage of R-LLGMN. The next subsection briefly describes the learning algorithm for the NN.

2.3. *Learning algorithm*

A set of input vector streams $\tilde{\mathbf{x}}^{(n)} = (\mathbf{x}(1)^{(n)}, \mathbf{x}(2)^{(n)}, \ldots, \mathbf{x}(T_n)^{(n)})$ $(n = 1, \ldots, N)$ and the teacher vector $\mathbf{T}^{(n)} = (T_1^{(n)}, \ldots, T_c^{(n)}, \ldots, T_C^{(n)})^{\mathrm{T}}$ are given for the learning of the R-LLGMN. We assume that the network acquires the characteristics of the data through learning if, for all the streams, the last output of stream $\tilde{\mathbf{x}}^{(n)}$, namely $^{(5)}O^c(T_n)$ $(c = 1, \ldots, C)$, is close enough to the teacher signal $\mathbf{T}^{(n)}$. The energy function for the network is defined as

$$J = \sum_{n=1}^{N} J_n = -\sum_{n=1}^{N} \sum_{c=1}^{C} T_c^{(n)} \log {}^{(5)}O^c(T_n). \tag{14}$$

The learning process attempts to minimize J, that is, to maximize the likelihood that each teacher vector $\mathbf{T}^{(n)}$ is obtained for the input stream $\tilde{\mathbf{x}}^{(n)}$.

The weight modification $\Delta w_{k',k,m,h}^c$ for $w_{k',k,m,h}^c$ is defined as

$$\Delta w_{k',k,m,h}^c = -\eta \sum_{n=1}^{N} \frac{\partial J_n}{\partial w_{k',k,m,h}^c} \tag{15}$$

in a collective learning scheme, where $\eta > 0$ is the learning rate. Due to the recurrent connection in R-LLGMN, this paper uses the back-propagation through time (BPTT) algorithm (Werbos, 1990). It is supposed that the error gradient within a stream is accumulated and weight modifications are only computed at the end of each stream; the error is then propagated backward to the beginning of the stream. Using the chain rule for the stream

$\tilde{\mathbf{x}}^{(n)}$, $\frac{\partial J_n}{\partial w^c_{k',k,m,h}}$ in (15) can be expanded as follows:

$$\begin{aligned}
\frac{\partial J_n}{\partial w^c_{k',k,m,h}} &= -\sum_{t=0}^{T_n-1}\sum_{c'=1}^{C}\sum_{k''=1}^{K_{c'}} {}^{(n)}\Delta^{c'}_{k''}(t)\frac{\partial {}^{(4)}O^{c'}_{k''}(T_n-t)}{\partial {}^{(4)}I^{c}_{k}(T_n-t)} \\
&\quad\times \frac{\partial {}^{(4)}I^{c}_{k}(T_n-t)}{\partial {}^{(3)}O^{c}_{k',k}(T_n-t)}\frac{\partial {}^{(3)}O^{c}_{k',k}(T_n-t)}{\partial {}^{(3)}I^{c}_{k',k}(T_n-t)}\frac{\partial {}^{(3)}I^{c}_{k',k}(T_n-t)}{\partial {}^{(2)}O^{c}_{k',k,m}(T_n-t)} \\
&\quad\times \frac{\partial {}^{(2)}O^{c}_{k',k,m}(T_n-t)}{\partial {}^{(2)}I^{c}_{k',k,m}(T_n-t)}\frac{\partial {}^{(2)}I^{c}_{k',k,m}(T_n-t)}{\partial w^c_{k',k,m,h}} \\
&= -\sum_{t=0}^{T_n-1}\sum_{c'=1}^{C}\sum_{k''=1}^{K_{c'}} {}^{(n)}\Delta^{c'}_{k''}(t)\big(\Gamma_{(c',k''),(c,k)} - {}^{(4)}O^{c'}_{k''}(T_n-t)\big) \\
&\quad\times \frac{{}^{(4)}O^{c'}_{k''}(T_n-t)}{{}^{(4)}I^{c'}_{k''}(T_n-t)}{}^{(4)}O^{c}_{k'}(T_n-t-1){}^{(2)}O^{c}_{k',k,m}(T_n-t)X_h(T_n-t),
\end{aligned} \tag{16}$$

where ${}^{(n)}\Delta^{c'}_{k''}(t)$ is defined as the partial differentiation of J_n to ${}^{(4)}O^{c'}_{k''}(T_n-t)$,

$${}^{(n)}\Delta^{c'}_{k''}(t) = \frac{\partial J_n}{\partial {}^{(4)}O^{c'}_{k''}(T_n-t)}, \tag{17}$$

and $\Gamma_{(c',k''),(c,k)}$ is defined as

$$\Gamma_{(c',k''),(c,k)} = \begin{cases} 1 & (c'=c; k''=k) \\ 0 & (\text{otherwise}) \end{cases}. \tag{18}$$

(17) can be expanded as follows:

$$\begin{aligned}
{}^{(n)}\Delta^{c'}_{k''}(0) &= \frac{\partial\big(T^{(n)}_{c'}\log {}^{(5)}O^{c'}(T_n)\big)}{\partial {}^{(5)}O^{c'}(T_n)}\frac{\partial {}^{(5)}O^{c'}(T_n)}{\partial {}^{(4)}O^{c'}_{k''}(T_n)} \\
&= \frac{T^{(n)}_{c'}}{{}^{(5)}O^{c'}(T_n)},
\end{aligned} \tag{19}$$

$$\begin{aligned}
{}^{(n)}\Delta^{c'}_{k''}(t+1) &= \sum_{c''=1}^{C}\sum_{k'''=1}^{K_{c''}} {}^{(n)}\Delta^{c''}_{k'''}(t)\sum_{k''''=1}^{K_{c''}}\frac{\partial {}^{(4)}O^{c''}_{k'''}(T_n-t)}{\partial {}^{(4)}I^{c'}_{k''''}(T_n-t)} \\
&\quad\times \frac{\partial {}^{(4)}I^{c'}_{k''''}(T_n-t)}{\partial {}^{(3)}O^{c'}_{k'',k''''}(T_n-t)}\frac{\partial {}^{(3)}O^{c'}_{k'',k''''}(T_n-t)}{\partial {}^{(4)}O^{c'}_{k''}(T_n-(t+1))} \\
&= \sum_{c''=1}^{C}\sum_{k'''=1}^{K_{c''}} {}^{(n)}\Delta^{c''}_{k'''}(t)\sum_{k''''=1}^{K_{c''}}\big(\Gamma_{(c'',k'''),(c',k'''')} - {}^{(4)}O^{c''}_{k'''}(T_n-t)\big) \\
&\quad\times \frac{{}^{(4)}O^{c''}_{k'''}(T_n-t)}{{}^{(4)}I^{c''}_{k'''}(T_n-t)}{}^{(3)}I^{c'}_{k'',k''''}(T_n-t).
\end{aligned} \tag{20}$$

In this paper, the dynamics of the terminal attractor (TA) (Zak, 1988) is incorporated into the learning rule in order to regulate the convergence time of the learning. The terminal attractor is based on the concept that the Lipschitz conditions are violated at the equilibrium point. The network learning converges to the equilibrium point, that is, the global minimum or local minima, in a finite specified time (Fukuda et al., 1995).

3. EMG discrimination method

The structure of the proposed discrimination method is shown in figure 3. This method consists of three parts in sequence: (1) EMG signal processing, (2) Recurrent neural network, and (3) Discrimination rule. The discriminated motion is used as the *command* for prosthetic control.

3.1. EMG signal processing

The EMG signals are processed to extract the feature patterns. The EMG signals measured from L pairs of electrodes are rectified and filtered by a second-order Butterworth filter (cut-off frequency: 1 [Hz]). They are then digitized by an A/D converter with a sampling frequency of f_d [Hz]. Sampled data is defined as $\text{EMG}_l(t)(l = 1, \ldots, L)$ and is normalized to make the sum of L channels equal to 1:

$$x_l(t) = \frac{\text{EMG}_l(t) - \text{EMG}_l^{\text{st}}}{\sum_{l'=1}^{L}\left(\text{EMG}_{l'}(t) - \text{EMG}_{l'}^{\text{st}}\right)}\ (l = 1, \ldots, L), \tag{21}$$

where EMG_l^{st} is the mean value of $\text{EMG}_l(t)$ measured while the arm is relaxed. The feature vector $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_L(t)]$ is used for the input of R-LLGMN. In this paper, we assumed that the amplitude level of the EMG signal is changed in proportion to muscle force. Force information $F_{\text{EMG}}(t)$ for the input vector $\mathbf{x}(t)$ is defined as follows:

$$F_{\text{EMG}}(t) = \frac{1}{L}\sum_{l=1}^{L}\frac{\text{EMG}_l(t) - \text{EMG}_l^{\text{st}}}{\text{EMG}_l^{\text{max}} - \text{EMG}_l^{\text{st}}}, \tag{22}$$

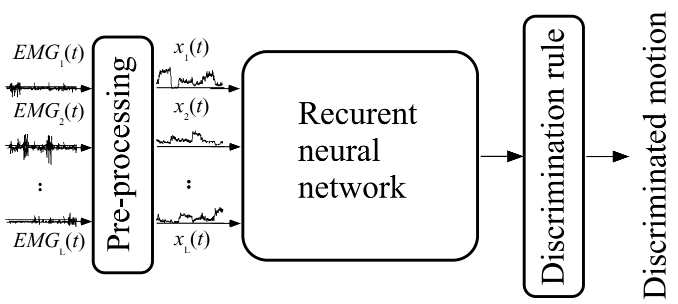

Figure 3. Structure of the proposed method.

where $\mathrm{EMG}_l^{\max}$ is the mean value of $\mathrm{EMG}_l(t)$ measured while maintaining the maximum arm voluntary contraction.

3.2. *Recurrent neural network*

R-LLGMN described in Section 2.2 is employed for pattern discrimination. Using samples labeled with the corresponding motions, R-LLGMN learns the non-linear mapping between the EMG patterns and the forearm motions. Given an EMG feature stream $\mathbf{x}(t)(t = 1, \ldots, T)$, the output ${}^{(5)}O^c(T)$ $(c = 1, \ldots, C)$ presents the a posteriori probability of each discriminating motion.

3.3. *Discrimination rule*

In order to recognize whether the motion has really occurred or not, the force information $F_{\mathrm{EMG}}(t)$ is compared with the prefixed motion appearance threshold M_d. The motion is considered to have occurred if $F_{\mathrm{EMG}}(t)$ exceeds M_d. The entropy of the R-LLGMN's outputs is also calculated to present the risk of misdiscrimination. The entropy is defined as

$$H(t) = -\sum_{c=1}^{C} {}^{5}O^c(t) \log_2 {}^{5}O^c(t). \tag{23}$$

If the entropy $H(t)$ is less than the discrimination threshold H_d, specific motion whose probability is the largest is determined according to the Bayes decision rule. If not, the determination is suspended.

4. Experiments

4.1. *Experimental conditions*

EMG pattern discrimination experiments were conducted to examine the performance of the proposed method with five subjects (amputee subjects A and B, and normal subjects C, D, and E).

Subject A (male) lost his forearm about 3 cm from the left wrist. EMG signals were measured from six pairs of electrodes ($L = 6$), at a sampling frequency of $f_d = 60$ [Hz]. The electrodes were attached to his forearm and upper arm (Flexor Carpi Radialis (FCR), Extensor Carpi Ulnaris (ECU), Flexor Carpi Ulnaris (FCU), Biceps Brachii (BB), Triceps Brachii (TB); two pairs on FCR and one pair on the others). The subject was asked to perform six motions ($C = 6$) in the order of flexion, extension, supination, pronation, hand grasping and hand opening, continuously for six-second periods.

Subject B (male) lost his right hand about 15 cm from the wrist. EMG signals were measured from eight pairs of electrodes attached to his forearm and upper arm ($L = 8$), and f_d was set at 100 [Hz]. EMG signals during eight motions ($C = 8$: flexion, extension, supination, pronation, hand grasping, hand opening, co-contraction of wrist joint and co-contraction of finger part) for 20-second periods were recorded continuously.

Subjects C (male), D (male) and E (male) are all normal-limbed, and six pairs of electrodes ($L = 6$) were attached in the same way as for subject A. EMG signals were measured (sampling frequency: 100 [Hz]) for 22 seconds, and seven motions ($C = 7$) were performed in the order of hand grasping, hand opening, extension, flexion, pronation, supination and co-contraction of finger part.

In the learning process of R-LLGMN, we used 20 EMG patterns extracted from the EMG signals for each motion and teacher signals consisting of $C \times 20$ patterns. According to the previous researches on EMG pattern discrimination, (Tsuji et al., 1993, 2001; Fukuda et al., 2000), the determination threshold H_d was set to 0.5, and the motion appearance threshold M_d to 0.2.

4.2. Results

An example of the discrimination result of subject A is shown in figure 4. In this figure, six channels of the input EMG signals, the force information $F_{\text{EMG}}(t)$, the entropy $H(t)$ and

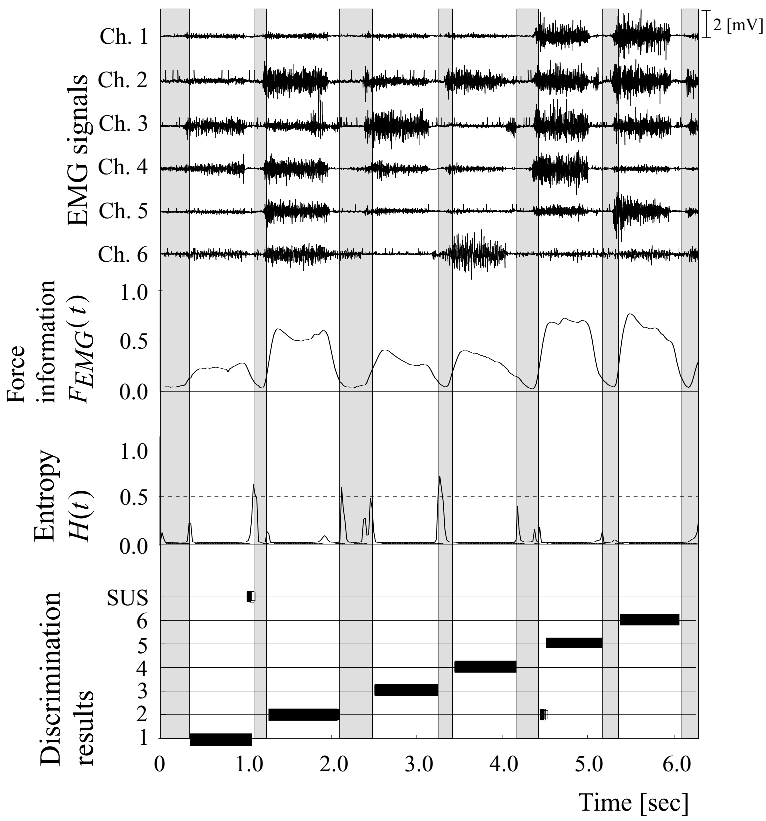

Figure 4. An example of the discrimination result.

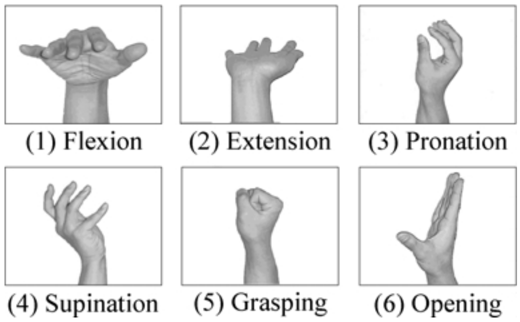

Figure 5. Motions.

the discrimination results are plotted. The labels of the vertical axis in the discrimination results correspond to the motions shown in figure 5, where SUS means that the determination was suspended. The gray areas indicate that no motion was determined because the force information is less than M_d. Incorrect determination was eliminated using the entropy. Figure 4 shows that the proposed method achieves high discrimination performance, even with non-stationary EMG signals during continuous motion.

We also investigated the accuracy of the discrimination results for five subjects (A, B, C, D and E). To verify the discrimination performance of the proposed method, we compared R-LLGMN with LLGMN and BPNN in experiments. The same preprocessing method and discrimination rule (see Sections 3.1 and 3.3) were applied to the experiments using LLGMN and BPNN. The LLGMN is a three-layer probabilistic NN based on GMM (Tsuji et al., 1999), and is a kind of feedforward neural networks with no recurrent connection. The number of units in the input layer of LLGMN was equal to the dimension of input signal (L). Units in the hidden layer corresponded to the Gaussian components in GMM, the number of which was set in the same manner as the R-LLGMN. The output layer included C units, and each unit gave the a posteriori probability for the input pattern. On the other hand, BPNN had four layers (two hidden layers) without any recurrent connection, and the units of the layers were set at 8, 10, 10 and 8. Each output of BPNN corresponds to a motion, so that it was normalized to make the sum of all outputs equal 1.0 to compare with R-LLGMN and LLGMN. The learning procedure of BPNN continued until the sum of the square error was less than 0.01, where the learning rate was 0.01. However, if the sum of the square error after 50,000 iterations was still not less than 0.01, the learning procedure was stopped. In all three methods, 10 different sets of initial weights (all randomized between [0, 1]) were used.

The mean values and the standard deviations of the discrimination rates are shown in Table 1. It can be seen that R-LLGMN achieved the best discrimination rate among all three methods, and had the smallest standard deviation.

Finally, the discrimination results were examined by altering the experimental conditions such as the length of sample data. Experiments were performed using various lengths of sample data. For each sample data, R-LLGMN was trained with 10 different sets of initial weights, which were randomly chosen in [0, 1]. The mean values of the discrimination rates

Table 1. Discrimination results of five subjects.

Subject	A	B	C	D	E
R-LLGMN	99.06	89.32	93.04	93.49	92.75
Mean ± SD (%)	±0.00	±0.37	±0.11	±0.00	±0.00
LLGMN	94.00	82.83	88.50	88.67	89.26
Mean ± SD (%)	±5.50	±0.00	±0.04	±0.15	±0.14
BPNN	73.41	46.52	44.20	69.79	69.17
Mean ± SD (%)	±7.86	±12.3	±10.4	±9.97	±7.00

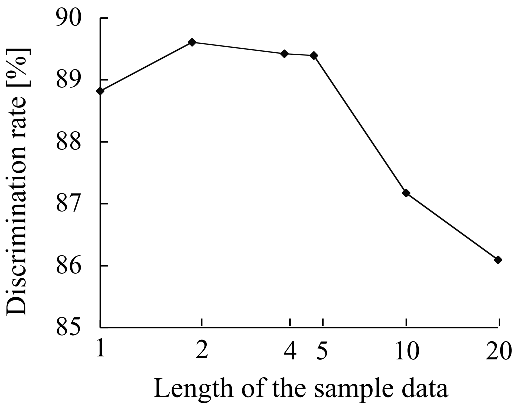

Figure 6. Discrimination rates for various data lengths (Subject B).

for each length are shown in figure 6, where the standard deviations are all very small, close to 0. It can be seen from figure 6 that the discrimination rate maintains a high level when the sample data is of an appropriate length (T). However, if $T > 5$, it is too long to train R-LLGMN, and the discrimination rate tends to deteriorate because R-LLGMN, which was trained using the long-length sample data, failed to discriminate the switching of motions.

5. Conclusions

This paper proposes a new EMG discrimination method based on a recurrent log-linearized Gaussian mixture network (R-LLGMN) for prosthetic control. Because of the recurrent connections between the third and the fourth layers in the R-LLGMN, the temporal information in the EMG signal can be used for the pattern discrimination.

To examine the discrimination capability and the accuracy of the proposed method, we conducted EMG pattern discrimination experiments with five subjects. In the experiments, the proposed method achieved high discrimination performance for varying EMG signals, and its discrimination results are superior to those of LLGMN and BPNN.

In future research, we would like to develop a new pre-processing method for the EMG signal. Discrimination performance could possibly be improved using a combination of a new pre-processing method and the R-LLGMN.

Acknowledgments

This work was partly supported by the Industrial Technology Research Grant Program in 2001 from the New Energy and Industrial Technology Development Organization (NEDO) of Japan.

References

Fukuda, O., Tsuji, T., and Kaneko, M. (1995). Pattern Classification of EEG Signals Using a Log-Linearized Gaussian Mixture Neural Networks. In *Proc. IEEE International Conf. Neural Networks 1995*, Vol. V (pp. 1113–1115). Perth, Australia.

Fukuda, O., Tsuji, T., and Kaneko, M. (2000). A Human Supporting Manipulator Based on Manual Control Using EMG Signals. *Journal of the Robotics Society of Japan*, 18(3), 387–394 (in Japanese).

Graupe, D., Magnussen, J., and Beex, A.A.M. (1978). A Microprocessor System for Multifunctional Control of Upper Limb Prostheses via Myo Electric Signal Identification. *IEEE Trans. on Automatic Control*, 23(4), 538–544.

Hiraiwa, A., Shimohara, K., and Tokunaga, Y. (1989). EMG Pattern Analysis and Classification by Neural Network. In *Proc. of IEEE International Conf. on Syst., Man and Cybern.* (pp. 1113–1115).

Hiraiwa, A., Uchida, U., and Shimohara, K. (1992). EMG/EEG Pattern Recognition by Neural Networks. *Proc. of the Eleventh European Meeting on Cybernetics and Systems Research* (pp. 1383–1390).

Huang, H.P. and Chen, C.Y. (1999). Development of a Myoelectric Discrimination System for a Multi-Degree Prosthetic Hand. In *Proc. of the 1999 IEEE International Conf. on Robotics and Automation* (pp. 2392–2397). Detroit, USA.

Kelly, M.F., Parker, P.A., and Scott, R.N. (1990). The Application of Neural Networks to Myoelectric Signal Analysis: A Preliminary Study. *IEEE Trans. on Biomedical Engineering*, 37(3), 221–230.

Koike, Y. and Kawato, M. (1994). Estimation of Arm Posture in 3D-Space from Surface EMG Signals Using a Neural Network Model. *Transactions Institute of Electronics, Information and Communication Engineers*, J77-D(4), 368–375 (in Japanese).

Rabiner, L.R. (1989). A Tutorial on Hidden Markov Model and Selected Applications in Speech Recognition. *Proc. of the IEEE*, 77(2), 257–286.

Tsuji, T., Bu, N., Fukuda, O., and Kaneko, M. (2003). A Recurrent Log-Linearized Gaussian Mixture Network, *IEEE Trans. Neural Networks*, 14(2), 304–316.

Tsuji, T., Bu, N., Murakami, M., and Kaneko, M. (2001). Pattern Discrimination of Raw EMG Signals Using a New Recurrent Neural Network. In *Proc. of IMEKO/SICE/IEEE the Int. Symp. on Measurement, Analysis and Modeling of Human Functions* (pp. 96–101). Sapporo, Japan.

Tsuji, T., Fukuda, O., Ichinobe, H., and Kaneko, M. (1999). A Log-Linearized Gaussian Mixture Network and its Application to EEG Pattern Classification. *IEEE Trans. Systems, Man, and Cybernetics-Part C: Applications and Reviews*, 29(1), 60–72.

Tsuji, T., Ichinobe, H., Ito, K., and Nagamachi, M. (1993). Discrimination of Forearm Motions from EMG Signals by Error Back Propagation Typed Neural Network Using Entropy. *Transactions of the Society of Instrument and Control Engineers*, 29(10), 1213–1220 (in Japanese).

Tsuji, T., Ito, K., and Nagamachi, M. (1987). A Limb-Function Discrimination Method Using EMG Signals for the Control of Multifunctional Powered Prostheses. *Transactions Institute of Electronics, Information and Communication Engineers*, J70-D(1), 207–215 (in Japanese).

Tsuji, T., Mori, D., and Ito, K. (1992). Motion Discrimination Method from EMG Signals Using Statistically Structured Neural Networks. *The Transactions of The Institute of Electrical Engineers of Japan*, 112-C(8), 465–473 (in Japanese).

Werbos, P.J. (1990). Backpropagation Through Time: What it Does and How to do it. *Proc. of the IEEE*, 78(10), 1550–1560.

Zak, M. (1988). Terminal Attractors for Addressable Memory in Neural Networks. *Physics Letters A*, 133(1/2), 18–22.