The Important Frequency Band Selection and Feature Vecotor Extraction System by an Evolutional Method

Yuuki Yazama*, Yasue Mitsukura**, Minoru Fukumi* and Norio Akamatsu*

*University of Tokushima, 2-1 Minami-josanjima Tokushima JAPAN
(Email:{rabbit, fukumi, akamatsu}@is.tokushima-u.ac.jp)
**Okayama University, 3-1-1 Tsushima-Naka Okayama JAPAN
(Email:mitsue@cc.okayama-u.ac.jp)

Abstract: In this paper, we propose the method to extract the important frequency bands from the EMG signal, and for generation of feature vector using the important frequency bands. The EMG signal is measured with 4 sensor and is recorded as 4 channel's time series data. The same frequency bands from 4 channel's frequency components are selected as the important frequency bands. The feature vector is calculated by the function formed using the combination of selected same important frequency bands. The EMG signals acquired from seven wrist motion type are recognized by changing into the feature vector formed. Then, the extraction and generation is performed by using the double combination of the genetic algorithm (GA) and the neural network (NN). Finally, in order to illustrate the effectiveness of the proposed method, computer simulations are done.

Keywords: EMG signal, neural network, genetic algorithm, extraction of feature, feature vector

1. Introduction

Electromyograph (EMG) is an electrical recording of muscle activity which is measured from the surface of the skin. It is possible to perform control of a manipulator or prosthetic hand by utilizing EMG as an incoming signal [1]-[3]. A facial expression and gesture is used as a natural interface of man, and a machine and a computer. However systems using EMG which can be used as interface more easily have studied since 1960s. The myoelectric upper limb prosthesis using EMG has a significant role in the field of rehabilitation medicine or welfare as an equipment which realizes natural motions as man's hand by engineering progress [4]. Furthermore personal digital assistant (PDA) apparatuses, for example cellular phone, have spread quickly. The portability of the PDA, miniaturization and weight saving are developing the technology of the PDA and radio-communications standards, such as bluetooth. If we use this technology, we can access to various apparatus using one operation apparatus to which the operation used frequently is collected. Then, it is thought that a wristwatch-type or a ring-type are desirable as for comprehensive operation equipment by EMG signal. In addition, a feature vector in a few dimension is necessary in order to aim at improvement of recognition accuracy. A neural network (NN) which has on-line type separation capability is useful for classification of EMG patterns.

Generally, the biomedical signal has difference among the individuals but reproducibility is comparatively high in the same person. This reason originates from the difference of muscular amount and transmission characteristic to the skin. Therefore, the technique for acquiring the individual characteristic has been a main. In this research, a back-propagation trained network is used, which can obtain a nonlinear map by learning. It can be seen that NN is suitable for modeling feature like the living body signal with nonlinear property. Moreover, to improve the recognition accuracy, the feature extraction becomes an important problem. Especially, the extraction of the feature in the pattern recognition is one of the important problems, and decides the performance in the recognition system. As the main feature extraction method, there are the principal component analysis to convert into the feature to represent at feature space and the feature selection method to extract only the feature to contribute to the improvement of recognition accuracy.

In this paper, we propose the recognition system of EMG signal that apply the combination of NN and GA by two stage. First of all, the extraction of important same frequency bands to each signal is performed by using the first combination of GA and NN (GANN). Then, the function that change the frequency spectrum into a feature vector is obtained by using the GANN. Therefore, the NN is used for modeling the function changed into the feature vector. The double structured GA is applied to each subject, and the feature vector specialized in the individual is extracted. Finally, we carry out the motion recognition experiment of the EMG signal by using the feature vector and show the effectiveness of the proposed method.

2. EMG signal

We focused on 7 operations which are considered as basic wrist motion types, neutral, top, bottom, left, right, outward rotation and inward rotation. The EMG patterns are recorded under the wrist motion. These operations are thought as standard motions and reflect a subject's simple operation.

Data extraction is preformed for one male subject and 30 sets are measured, each of which is composed of 7 operations. The EMG signals are measured by a surface electrode. Moreover, the surface electrode of dry type (electrolysis cream was not used) is adopted because of a subject's displeasure and a practical application as PDA operation machines. The EMG pattern which is amplified electrically the feeble signal emitted from a living body, and is obtained as time series data. 4 electrodes are placed around the wrist (see Figure.1). The

This paper is supported by my funding agencies.

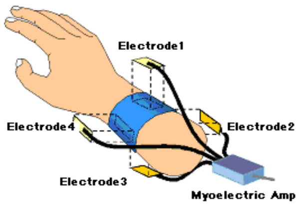

Fig. 1. Wrist with 4 electrodes

EMG patterns are measured under sampling frequencies of 20.48kHz and for 2.048 seconds in all electrodes.
The Fast Fourier Transform (FFT) is used for the time series data of the EMG patterns. The feature vector is generated as frequency distribution of the EMG pattern. The normalization is performed to 4ch signals at the maxinum of a power spectrum of the signal of 4ch, and frequency band to be used is from 40Hz to 1000Hz.

3. The construction of EMG recognition system

This chapter describes the composition of the proposal system. The double GANN are carried out for computing the feature vector. The operation recognition is enabled by a little feature dimensions. Next, the important frequency band is extracted from an superior chromosome obtained by the first GANN. And, the function is modeled to compute the feature vector using the important frequency bands by the sencond GANN. Finally, the recognition experiment of the EMG signal is performed by NN and the obtained function (see Figure.2).

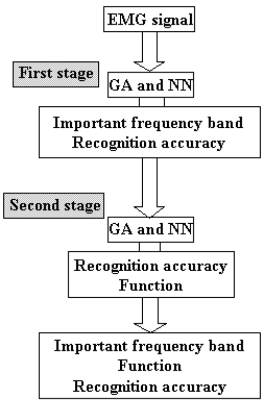

Fig. 2. The process of EMG recognition system

GA in the first stage has 1 dimension chromosome, and GA in the secondly stage has 2 dimension chromosome. NN is consistently learned by using the backpropagation method (BP). As for the composition of NN, the number of input layers is enable to change if necessary. Besides, the number of hidden layer is 10, and the number of output layer is 7 as the number of identification wrist motion type.

4. The detail of EMG recognition system

This method to recognize the EMG signal is a proposal of the system that applies the double combination of GA and NN. Details to each stage are described as follows.
In this research, the backpropagation trained network is used, which can obtain a nonlinear map by learning. It can be seen that NN is suitable for modeling feature like the biomedical signal with nonlinear property. NN has a three-layered structure. And GA is a method of emulating the evolution process in a life. GA can be obtained a better solution by changing two or more solutions being inherited. This is one of the optimization methods to be effective to a global search and is used for selecting the important frequency bands and forming the function.

4.1. The structure of neural Network

NN is the suitable method for modeling the feature of the biomedical signal that the nonlinearity is strong. Then, the backpropagation method is applied to learn for NN of three layer structure. The structure applied by one stage and two stages is basically the same excluding the number of input layers. The Table.1 is shown the setup parameter in NN.

Table 1. The setup of parameter in NN

The number of input layers	variable
The number of hidden layers	10
The number of output layers	7
The number of iterations	30000
THe learning parameter	0.05

4.2. The selection of important frequency bands

It is thought that unnecessary frequency bands is included in the frequency bands for the recognition. In general, the learning of NN and recognizing the wrist motion by using unnecessary feature cause the decrease of the recognition accuracy. Then, the important frequency bands is first selected by GANN in the first stage.

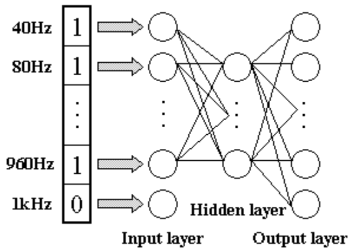

Fig. 3. The outline of the first GANN

It explains the chromosome used with GA in the first stage. The chromosome is binary coding of one dimension. Each gene corresponds to each frequency band. The frequency band corresponding to code “1 ” is used for learning and recognition. However, the frequency band corresponding to code “0 ”. is not used for learning and recognition. In addition, each gene of the chromosome corresponds to the input layer of NN by the one to one. The evaluation of the chromosome is done by the recognition accuracy and the data deletion rate. The fitness value is calculated as follow :

$$F_{1i} = Rec_i \times Gene_i \tag{1}$$

where Rec_i means the recognition accuracy and $Gene_i$ means the reduction rate of frequency bends. It is thought that unnecessary data can be deleted as maintaining and improving the recognition accuracy.
The EMG signal is obtained by 4ch signals. The same chromosome is applied to 4ch EMG frequency components. In a word, the important same frequency bands is selected from 4ch EMG frequency components. I think that the same frequency band has influenced each other. The second GANN obtain the function which the feature vector calculated using the same important frequency band. In addition, it is thought that the important frequency not to depend on the acquisition part band is selected. The genetic operator is one point crossover, mutation, roulette selection and elite strategy method. Table.2 is shown a setup parameter of GA in the first stage.

Table 2. The setup of parameter in the first GA

The number of generations	50
The number of chromosomes	50
The bit length	25
The rate of elite	012
The rate of crossover	0.6
The rate of mutant	0.28

4.3. The modeling of function

In second stage of GANN, it forms the function which generate the feature vector using only the selected frequency bands obtained from the result of the first GANN. The derived function is composed of the addition and subtraction multiplication and division calculation. Moreover, as for division calculation, the arrangement of the molecule and the denominator also is searched by GA. This function generates a new feature vector by using the selected frequency. And, the evaluation of the function is evaluated at the recognition accuracy.
First, the chromosome structure of second GA is used 2-dimension structure(see Figure.4). The chromosome structure is that the line shows the frequency band and the row shows the function. The frequency bands selected by the first GANN are used. The row of chromosome structure shows one function by the unit of 3 bit. The EMG signal is 4ch signals. And, the first 3 bit from the left shows the function that the frequency band of ch1 and ch2 combined. The next 3 bit shows the function that the frequency band of ch1 and ch3 combined. 3 bit to form the function is assumed to be a minimum composition. The table.3 shows the calculation type that the 3 bit expresses.

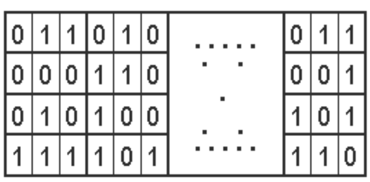

Fig. 4. 2-dimension chromosome

Table 3. The operational expression using the gene

addition	0	0	$*$	$a+b$
subtraction	0	1	$*$	$a-b$
multiplication	1	0	$*$	$a \times b$
division 1	1	1	1	a/b
division 2	1	1	0	b/a

n in Table.3 is shown the number of frequency band selected. And The third bit of a minimum composition is used at only division calculation and shows the arrangement of the molecule and the denominator. The variable a and b is shown the frequency spectrum used. $*$ in Table.3 is shown the arbitrary bit.
The feature vector is calculated at each combination of 2ch signal in 4ch signal by the function. When one frequency band is selected, $6({}_4C_2)$ kinds of feature vectors are generated. The number of generated feature vectors is the numbers six times the number of frequency bands chosen by the first GANN.
Next, it explains the genetic operator used. The genetic operator uses the same method of the first GANN. However, the process of crossover and mutation is performed by specifying the range of crossover and mutation (see Figure.5 and Figure.6). The range of processing is decided at random. The Table.4 shows the parameter of the second GA.

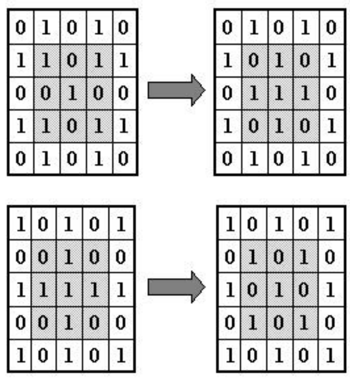

Fig. 5. Crossover operation

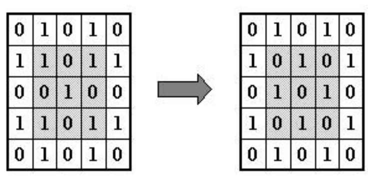

Fig. 6. Mutation operation

The second GANN is processed to 5th chromosomes with the highest recognition accuracy obtained among all chromosomes in the first GANN. And the feature vector is generated with the function which uses the important frequency bands of each subject. A final simulation is performed only by NN using the feature vectors, and the effectiveness of this method is verified.

Table 4. The setup of parameters in the second GA

The number of generations	50
The number of chromosomes	50
The bit length	18
The rate of elite	01
The rate of crossover	0.7
The rate of mutant	0.2

5. Computer simulation

The computer simulation is performed by using the combination of GA and NN which is two stage type, and the EMG acquired from one subject. The collection EMG data for each subject is 30 sets of which the 1 set include seven operation. 10 sets are used for learning data, and other 20 sets are used for test data. These data sets are changed at random every generation. By doing so, I think that the high result with reliability can be gotten. Table.5 show a result of EMG recognition experiment acquired by using NN. The recognition accuracy shows the average recognition accuracy obtained from 30 times computer simulation. In addition maximum recognition accuracy and minimum recognition accuracy, which obtained among 30 times computer simulation, also shows. Then, that computer simulation is performed by using the important frequency bands extracted from the most superior chromosome selected by the first GANN.

Table 5. The recognition accuracy using the important frequency bands

	N	T	U	R
*	86.1	70.8	77.6	88.5
**	95.0/70.0	90.0/50.0	90.0/80.0	85.0/80.0
	L	RI	RO	Total
*	95.6	85.1	81.6	83.6
**	100.0/80.0	70.0/80.0	90.0/85.0	88.5/75.7

* : The average recognition accuracy

** : The recognition accuracy of maximum / minimum

Moreover, Table.6 shows the result of recognition experiment acquired by using the feature vector calculated by the function made by the second GANN.

Table 6. The recognition accuracy using the feature vectors

	N	T	U	R
*	89.6	75.2	91.7	89.3
**	100.0/75.0	85.0/55.0	95.0/95.0	95.0/75.0
	L	RI	RO	Total
*	94.2	86.4	81.1	86.8
**	100.0/85.0	95.0/85.0	95.0/75.0	95.0/77.8

* : The average recognition accuracy

** : The recognition accuracy of maximum / minimum

6. Consideration

To compare the obtained result, the recognition accuracy by the frequency bands from 40Hz to 1000Hz is shown in Table.7.

Table 7. The recognition accuracy using the frequency band from 40Hz to 1000Hz

	N	T	U	R
*	87.1	59.6	61.6	74.3
**	100.0/85.0	65.0/60.0	74.0/45.0	80.0/75.0
	L	RI	RO	Total
*	89.0	76.3	76.3	76.4
**	100.0/75.0	85.0/55.0	90.0/85.0	85.0/68.5

* : The average recognition accuracy

** : The recognition accuracy of maximum / minimum

Here, the reduction of 64% feature can be done by the extraction of the important frequency bands. Moreover, the recognition accuracy has improved by 7.2% by using 36 features. And, the number of final feature becomes $9 \times 6 = 54$ because the feature vectors are calculated by using 9 important frequency bands. As for the number of feature vectors, the reduction in 46% is done compared with the number of feature that uses from 40Hz to 1000Hz.

7. Conclusions

In this paper, we propose the EMG recognition system by using the two stage combination of GA and NN. Here, the combination of GA and NN perform the extraction of important frequency bands and the calculation of feature vectors. As a result, the average recognition accuracy of 10% can be improved and the 46% reduction of feature dimension can be performed.

As a future work, the reduction of the calculated feature vector is perform. I think that the recognition accuracy can be improved by using less feature dimension by doing so. Moreover, I think that the identification of wrist motion type can be carried out easily by the extraction of feature signal that becomes main when the wrist operates.

References

[1] D.Nishikawa, W Yu, H.Yokoi and Y.Kakazu. *n-Line Learning Method for EMG Prosthetic Hand Controlling.*, *IEICE*, D-II Vol J82, No.9, pp.1510-1519, September 1999, in japanese.

[2] D.Nishikawa, H.Yokoi, and Y.Kakazu. *Design of motion-recognizer using electoromyogram.*, In *Robotics and Mechatronics Division*, June 1998, in japanese.

[3] D.Nishikawa, W Yu, H.Yokoi, and Y.Kakazu. *On-Line Supervising Mechanism for Learning Data in Surface Electromyogram Motion Classifiers.*, In *IEICE*, D-II Vol J84, No.12, pp.2634-2643, December 2001.

[4] E.Lamounier, A.Soares, A.Andrade, and R.Carrijo. *A Virtual Prosthesis Control Based on Neural Networks for EMG pattern Classification*, proc.of IASTED, Int. Conf. of AI&SC, Banff, Canada, July 2002.