The Seventeenth International Symposium on Artificial Life and Robotics 2012 (AROB 17th '12),
B-Con Plaza, Beppu, Oita, Japan, January 19-21, 2012

The Technique of the Online Learning Method using SOM algorithm for nonlinear SVM

Hiroki Tamura[1], Takeshi Yoshimatsu, and Koichi Tanno

[1] Department of Electrical & Electronic Engineering, University of Miyazaki, 1-1 Gakuen-kibanadai-nishi, Miyazaki-city, Miyazaki, 889-2192, Japan
(Tel: 81-985-58-7409)

[1]htamura@cc.miyazaki-u.ac.jp

Abstract: The support vector machine (SVM) is known as one of the most influential and powerful tools for solving classification and regression problems, but the original SVM does not have an online learning technique. Therefore, many researchers have introduced online learning techniques to the SVM. In a previous article, we proposed an unsupervised online learning method using the technique of the self-organized map for the SVM. In another article, we proposed the midpoint validation method for an improved SVM. We test the performance of the SVM using a combination of the two techniques in this article. In addition, we compare its performance with the original hard margin SVM, and also experiment with our proposed method on surface electromyogram recognition problems with changes in the position of the electrode, and the numerical experiment recognition problem with changes in the time. These experiments showed that our proposed method gave a better performance than the SVM and corresponded to the changing data.

Keywords: Support vector machine, Online learning, Midpoint validation, Pattern classification problem, Surface electromyog ram

1 INTRODUCTION

The support vector machine (SVM) proposed by Cotes and Vapnik is one of the most influential and powerful tools for solving classification problems [1][2][3][4][5].

We studied surface electromyogram (abbr. s-EMG) recognition using the SVM. The purpose was the development of a human interface using an s-EMG. In this study, we considered problems such as the changes in the s-EMG pattern caused by muscle fatigue, and the position of the sensor doing the measurements. In a previous article [6] we proposed the online unsupervised learning method using the technique of a self-organized map for the SVM. Furthermore, the proposed method has a technique for the reconstruction of a SVM.

In addition, we are studying a SVM which is not limited to the recognition of s-EMG. In this study, we pay attention to the problem of the deflection of the separating hyperplane in the input space of a nonlinear SVM, and propose an improved method [7][8]. We call this method the midpoint validation method. This method assumes a midpoint between the classes of training data and an index of the deflection, and moves the separating hyperplane according to the index. This method also has a technique for the reconstruction of a SVM.

These two studies both achieved good results. However, we had not previously combined the two methods. Therefore, we tested the performance of the SVM using a combination of the two techniques. Moreover, we compared its performance with the original hard-margin SVM, Winner Take All model (our previous online method [10]), and also experimented with our proposed method on s-EMG recognition problems with changes in the position of the electrode, and the numerical experiment recognition problem with changes in the time.

2 PROPOSED METHOD

In this section, we introduce the SVM, our online learning method. We also propose a method that combines the online learning method with the midpoint validation method [7][8].

2.1 SVM

A SVM is a mechanical learning system that uses a hypothetical space of linear functions in a high-dimensional feature space. A nonlinear SVM is expressed by Eqs. 1–3. Here, we used the Gaussian kernel given in Eq. 1 as the kernel function, while the SVM decision function g(**x**) and the output of the SVM are as given in Eqs. 2 and 3.

$$K(\mathbf{x},\mathbf{x}_i)=\exp\left(\frac{-\|\mathbf{x}-\mathbf{x}_i\|^2}{2\delta^2}\right) \quad \textbf{(1)}$$

$$g(\mathbf{x})=\sum_{i=1}^{N} w_i K(\mathbf{x},\mathbf{x}_i)+b \quad \textbf{(2)}$$

$$O=sign(g(\mathbf{x})) \quad \textbf{(3)}$$

2.2 Online learning method

In this subsection, we introduce an unsupervised online learning method using the self-organizing map (abbr. SOM) algorithm for the SVM and a restructuring technique.

Let the input space be denoted by $\mathbf{x}_{in} \in$R. $\mathbf{x}_{in}$ ($in \notin \{i = 1, \ . \ . \ . \ , N\}$) is the input vector without the label. The training vectors are included in the kernel function x_i, with $i = 1, \ . \ . \ . \ , N$, which belongs to either of the two classes. Therefore, these are given the label $y_i \in \{-1, 1\}$. Each training vector has the same dimensions in input space. The flow of our online learning method is shown in Figs. 1 and 2. The parameter η is an update parameter. The idea of this rule is close to adaptive resonance theory. The parameter s is the Euclidean distance between $\mathbf{x}_{in}$ and each $\mathbf{x}_i$.

Step 1: To find the best mach of the input vector $\mathbf{x}_{in}$ with the training vectors $\mathbf{x}_i$, the Euclidean distance between $\mathbf{x}_{in}$ and each $\mathbf{x}_i$ is computed (Fig. 2a).

Step 2: The following processing (Steps 3-4) are not done to $\mathbf{x}_{in}$ when the label of $\mathbf{x}_{win}$ is not the same as the label of the output result of SVM of $\mathbf{x}_{in}$.

Step 3: To find the best mach of $\mathbf{x}_{win}$ with the training vectors $\mathbf{x}_j$, the Euclidean distance $\mathbf{x}_{win}$, and each $\mathbf{x}_j$ is computed (Fig. 2b).

Step 4: If d_w is condition of the rule of Eq. 4, each $\mathbf{x}_i$ is update according to the learning rule of Eqs. 5 and 6 (Figs. 2c and 2d).

Step 5: Steps 1-4 are done to all input vector.

Step 6: Proposed method repeats *Num* cycles these processing (Steps 1-5) using same inputs vector $\mathbf{x}_{in}$.

$$d_w \le d_O \tag{4}$$

$$\mathbf{x}_i^{new} = \mathbf{x}_i^{old} + \eta \cdot f(s)(\mathbf{x}_{in} - \mathbf{x}_i^{old}) \tag{5}$$

$$f(s) = \begin{cases} \exp(-(3s)^2) | y_{win} = y_i & (6.a) \\ 0 | y_{win} \neq y_i & (6.b) \end{cases}$$

$$s = \|\mathbf{x}_{in} - \mathbf{x}_i\| \tag{7}$$

Fig. 1. The flow of the online learning method.

If the SV changes after the update, the SVM is restructured using the updated training vectors. Even if the training vectors change using Steps 1–5, maximizing the margin of the SVM is retained in this restructuring process. In this paper, we first perform the online learning method. Next, we create a SVM with the updated training data. Finally, we perform the midpoint validation method [7][8].

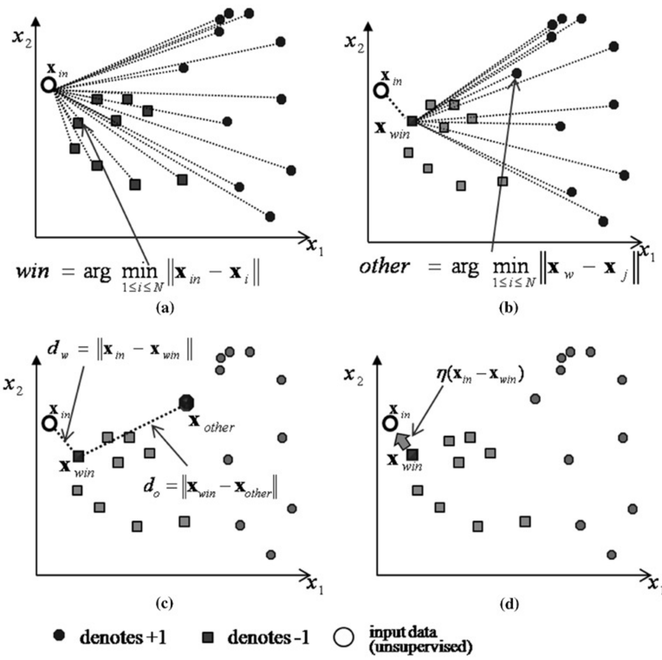

Fig. 2. The flow of the proposed method using SOM. The training vector x_{other} is selected as in (a) and (b). If d_w and d_o are the conditions of the rule in Eq. 4, as in (c), x_{win} is updated using the learning rule in Eq. 5 as in (d).

3 EXPERIMENTS

In this section, the system configuration for experiments on recognizing input data sets changes over time is explained. Next, the system configuration for experiments on recognizing forearm movements using s-EMG is explained. Finally, the results of the computer simulations are described. We compared the performance of the proposed method, Mexican Hat model, Winner Take All (abbr. WTA) model [10], the original hard-margin SVM, and *k*-NN method (form paper [10]). The Mexican Hat model is changed Eq.6.a and b of the proposed method to Eq.8. The WTA model is changed Eq.5 of the proposed method to Eq.9 and not use Eqs.6 and 7.

Mexican Hat model

$$f(s) = \begin{cases} \left(1 - \left(2 \cdot (12s)^2\right)\right) \exp(-(12s)^2) | y_{win} = y_i \\ 0 | y_{win} \neq y_i \end{cases} \tag{8}$$

WTA model [10]

$$\mathbf{x}_{win}^{new} = \mathbf{x}_{win}^{old} + \eta \cdot (\mathbf{x}_{in} - \mathbf{x}_{win}^{old}) \tag{9}$$

3.1 Experimental conditions on numerical

As time-varying input data sets, we created a spiral data sets vary in the two-dimensional space. This is two-class data sets. As class 1, 1000 data were calculated by adding the noise and according to Eqs.10 and 11. As class 2, shifted 180 [deg], 1000 data were created in the same way. In each class, data were divided into training and test data of 900 and 100. In addition, the test data was not chosen from the vicinity of the origin. Shifted 20 [deg] at an angle, 1000 data were created in the same way, in each class. And data were divided into learning and test data of 900 and 100, in each class. Repeat this process 9 times, made 9 data sets.

$$x = 0.1\theta\cos(8\pi\theta) \tag{10}$$

$$y = 0.1\theta\sin(8\pi\theta) \tag{11}$$

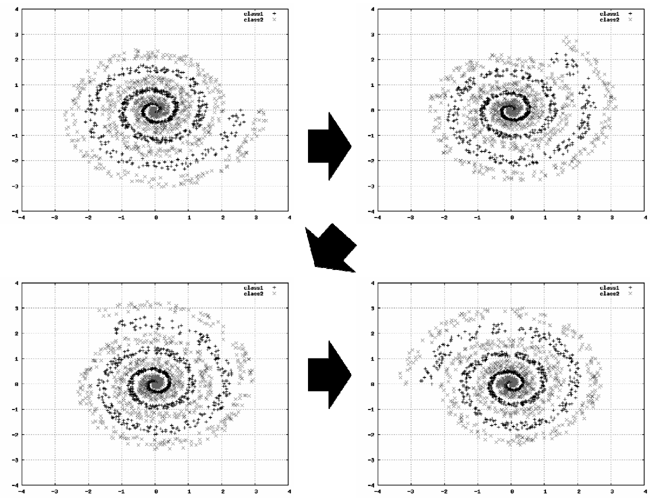

Fig. 3. The spiral data.

3.2 Experiment on numerical

We tested the effectiveness of the proposed method on the spiral recognition problem (Fig.3) that the quantity of the feature changes according to the angle. First, the SVM learn the relation between the coordinate and class from the training data (the training vectors). Next, rotate 20[deg] around the origin. Additional unsupervised learning data (the input vector) are obtained from each class. Last, the recognition rate is calculated from the test data for 100 repeats of each class. The experiments were tested nine times in total by rotating around the origin by 20[deg], 40[deg], 60[deg], 80[deg], 100[deg], 120[deg], 140[deg], 160[deg] and 180[deg]. The basis of the proposed method is the hard-margin SVM using Eq. 1. The Gaussian kernel parameters of the SVM were decided from an evaluation using the training data. Gaussian kernel parameter δ was 0.3. In this experiment, the value of parameter η was 0.1, and the value of *Num* was 1.

3.3 Experimental conditions on forearm muscles

The s-EMG of each movement pattern is measured with electrode sensors, and the quantity of the feature is extracted from the s-EMG. The quantity of the feature is given to the recognition machine as an input, and each movement pattern that generates s-EMG is presumed. The quantity of the feature uses the minimum–maximum values and integration values used by Tamura et al [9]. That article showed that the technique of min–max values and integration values is easier and better than fast Fourier transform processing. The sampling frequency of the measurement data was 1 KHz, and the sampling band was from 0 Hz to 500 Hz.

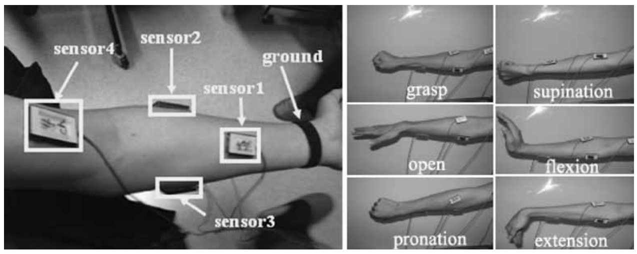

Fig. 4. Image of the forearm motion.

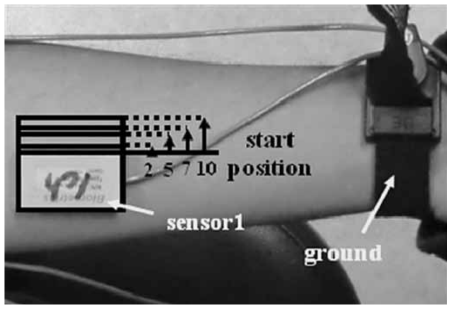

Fig. 5. S-EMG recognition problems with changes in the electrode position (2mm, 5mm, 7mm and 10mm).

3.4 Experiments on forearm muscles

We tested the effectiveness of the proposed method on the s-EMG recognition problem that the quantity of the feature changes according to the position of the electrode. The experimental subjects were three healthy men (T.Y, K.F and S.Y). The subjects sat on a chair. The recognition experiment of the six motion patterns was conducted using s-EMG obtained from four sensors set on the right arm (Fig. 4). Eight inputs were given to the identification machine. The experiments were conducted for 1 day. First, we acquired the training data from the s-EMG concerning the movements of the forearm. Then, the SVM learn the relation between the s-EMG and motion from the training data (the training vectors). Each movement is identified 60 times. Next, the electrode position (sensor 1) is moved 2 mm. Additional unsupervised learning data (the input vector: each motion is repeated 20 or 40 times) are obtained from each movement. Last, the recognition rate is calculated from the test data for 20 repeats of each movement. The measurements were tested four times in total by moving the electrode position by 2 mm, 5 mm, 7 mm, and 10 mm (Fig. 5). The basis of the proposed method is the hard-margin SVM using Eq. 1. The Gaussian kernel parameters of the SVM were decided from an evaluation using the training data. Subject T.Y was 0.7, K.F was 2.0, and S.Y was 0.9. In these experiments, the value of parameter η was 0.1 (on subject T.Y and S.Y) or 0.05 (on subject K.F), and the value of *Num* was 1.

3.5 Experimental result

The results for each simulation method are given in table 1 and 2. In table 1, we show the average success rate of experiment without midpoint validation. From table 1, the proposed method is better than WTA model on result of numeric and subject T.Y. However, Mexican Hat model is worse than the original SVM. Therefore, we consider three methods except the Mexican Hat model. In table 2, we show the average success rate of experiment with midpoint validation. From table 2, the proposed method is better than WTA model on result of numeric and subject T.Y. But, the proposed method is not effective on the subject S.Y. From these results, WTA model had better results, and the proposed method has specialty data and non-specialty data.

Table 1. Average success rate (%) of experiment without Midpoint validation method.

	Numerical	s-EMG		
		T.Y	K.F	S.Y
WTA model	56.3	80.0	**83.3**	**77.7**
Mexican Hat model	53.5	75.8	72.2	59.2
Proposed Method	**66.2**	**85.7**	81.7	72.2
original SVM	50.8	78.5	77.8	76.3

Table 2. Average success rate (%) of experiment with Midpoint validation method.

	Numerical	s-EMG		
		T.Y	K.F	S.Y
WTA model	54.6	81.7	**84.2**	**86.7**
Proposed Method	**68.1**	**86.3**	81.3	70.8
SVM+MV	48.1	77.0	80.8	81.3
k-NN method	-	80.8	82.1	77.7

4 CONCLUSIONS

In this paper, we have proposed a technique of the online learning method using SOM algorithm for nonlinear SVM. The experimental results show that the proposed method is effective in the data which changes uniformly like numerical problem. WTA model is effective in the data which changes irregularly. WTA model is suitable for the recognition problems of s-EMG, because to change irregularly in many cases. The merit of our proposed online method is that the proposed method can be used effectively by setting only one parameter η. The parameter η may be around 0.1. Therefore, the proposed method is useful. In future work, we will experiment on the effectiveness of the proposed method in other problems.

References

[1] Cotes C., Vapnik VN.: Support vector networks. Mach Learn 20(3):273–297 (1995)

[2] Burges CJC.: A tutorial on support vector machines for pattern recognition. Data Mining Knowledge Discovery 2(2) (1998)

[3] Joachims T.: Making large-scale support vector machine learning practical. In: Shoelkopf B, Burges C, Smola A (eds) Advances in kernel methods: support vector learning. pp 169–184 (1999)

[4] Mangasarian OL., Musicant DR.: Active support vector machines. Technical Report 00-04, Data Mining Institute, Computer Sciences Department, University of Wisconsin, Madison (2000)

[5] Mangasarian OL., Musicant DR.: Lagrange support vector machines. J Mach Learn Res 1:161–177 (2001)

[6] Tamura H., Yoshimatu T., Tanno K.: Support vector machines with online unsupervised learning method and its application to s-EMG recognition problems. NOLTA (2010)

[7] Tamura H., Tanno K.: Midpoint-validation method for support vector machine classification. IEICE Trans Inform Syst E91-D(7): 2095–2098 (2008)

[8] Yamashita S., Tamura H., Toyama T., et al.: The effectiveness of midpoint-validation method for support vector machines (in Japanese). Proceedings of Electronics, Information and Systems Conference, IEEJ, CD:GS12-5 (2010)

[9] Tamura H., Gotoh T., Okumura D., et al.: A study on the s-EMG pattern recognition using neural network. Int J Innovative Comput Inform Control 5(12B):4877–4884 (2009)

[10] Shingo Y., Takeshi Y., Hiroki T. and Koichi T.: A study of SVM using the Combination with Online Learning Method and Midpoint-Validation Method. Proc. of AROB2011 (2011)