Journal of University of Science and Technology Beijing
Vol.9(2002), No.1, p.74

Information

An Improved Minimum Distance Method Based on Artificial Neural Networks

Qing Li, Deling Zheng, Wenbo Meng, Yong Tang

Information Engineering School, University of Science and Technology Beijing, Beijing 100083, China
(Received 2001-03-28)

Abstract: MDM (minimum distance method) is a very popular algorithm in state recognition. But it has a presupposition, that is, the distance within one class must be shorter enough than the distance between classes. When this presupposition is not satisfied, the method is no longer valid. In order to overcome the shortcomings of MDM, an improved minimum distance method (IMDM) based on ANN (artificial neural networks) is presented. The simulation results demonstrate that IMDM has two advantages, that is, the rate of recognition is faster and the accuracy of recognition is higher compared with MDM.

Key words: state recognition; minimum distance method; artificial neural networks

[*This work was financially supported by the National Natural Science Foundation of China under the contract No.69372031.*]

1 Introduction of Minimum Distance Method and its Weakness [1]

MDM (minimum distance method) is a popular method for state recognition. It's usually operated as following steps.

(1) Determine the central vector W_i ($i=1, \cdots, c$, c is the total number of classes) of class M_i.

(2) Calculate the distance d_i ($i=1, \cdots, c$) between the vector X and the central vector W_i ($i=1, \cdots, c$) by the following equation:

$$d_i = \|X - W_i\| = \left[\sum_{k=1}^{n}(x_k - w_{ik})^2\right]^{1/2} \tag{1}$$

where $X=[x_1, x_2, \cdots, x_n]$ is the vector to be recognized and x_k ($k=1,\cdots, n$) is the characteristic value of vector X; $W_i=[w_{i1}, w_{i2}, \cdots, w_{in}]$ and w_{ik} ($k=1, \cdots, n$) is the characteristic value of the central vector W_i.

(3) Select the minimum distance d_j ($d_j=\min_i\|X-W_i\|$). Then vector X is labeled as class j.

When MDM is applied to state recognition, a presupposition, that is the distance within one class must be shorter enough than the distance between classes must be satisfied. For example, in **figure 1**, the data set M can be classified to 3 classes according to criteria A, which can be denoted as $M=\{M_1, M_2, M_3\}$. We can also see that M_1 and M_3 are composed of only one part and M_2 is composed of three parts. These three parts are called subclass and denoted as M_{21}, M_{22}, and M_{23}, i.e.

Corresponding author: Qing Li
E-mail: Li_Qing_2001@263.net

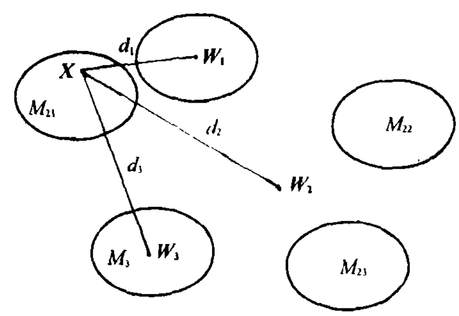

Figure 1 Sketch map of minimum distance method.

$M_2=\{M_{21}, M_{22}, M_{23}\}$. The subclass M_{21}, M_{22} and M_{23} can be clustered by criteria B. W_1, W_2 and W_3 are the central vectors of M_1, M_2 and M_3. If MDM is used to determine which class the vector X belongs to, then d_1, d_2 and d_3 need to be computed. It is obvious that d_1 is the shortest distances among d_1, d_2 and d_3. So we put the vector X to class 1. In fact, vector X belongs to class 2. The recognition result is wrong.

2 Improved Minimum Distance Method [2]

If we classify the data set M with criteria C ($C=A\cap B$), M can be grouped in 5 clusters as **figure 2** shows. In order to distinguish from the classes originally named, these 5 clusters are called new-classes and denoted as M_1', M_2', M_3', M_4' and M_5', i.e. $M=\{M_1', M_2', M_3', M_4', M_5'\}$. If the new-class satisfies the presupposition, MDM can be used to classify the data and will not make any mistakes. Then we can determine the central vector W_i' ($i=1, 2, \cdots, 5$) of M_i' $(1, 2, \cdots, 5)$ and calculate the distance $d_i'=\|X-W_i'\|$ ($i=1, 2, \cdots, 5$). For example, in figure 2, vector X can be put to new-class 2 because d_2' is the shortest distance among d_i' ($i=1, 2, \cdots, 5$).

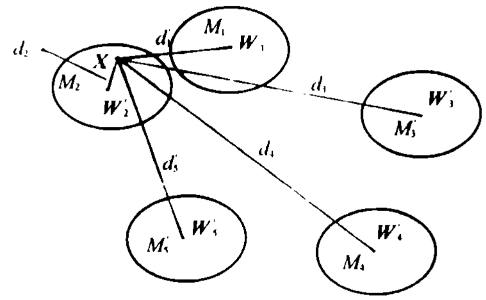

Figure 2 Sketch map of the improved minimum distance method.

Therefore we can put vector X to class 2 by use of the relationship between new-class and class as **table 1** shown. We can see that the recognition result is right. We call the algorithm above IMDM (improved minimum distance method).

Table 1 Relationship between new-class and class

new-class	1	2	3	4	5
class	1	2	2	2	3

When the improved minimum distance method is used for classification, there are two problems need to be solved.

(1) How to obtain criteria B so that the new-class can satisfy the presupposition.

(2) How to determine the central vector W'_i of new-class M'_i.

3 Realization of the Improved Minimum Distance Method Based on ANN (Artificial Neural Networks)

3.1 Model of neural networks [3]

The model of neural networks for state recognition is composeded of two layers as **figure 3** shows. In the first layer, every node is used to receive and normalize the characteristic value x_i $(i=0, \cdots, m-1)$ of vector X. Every node y_j $(j=0, \cdots, n)$ of the second layer denotes a new-class. The central vector of new-class j is expressed by weight vector $W_j=(w_{j0}, w_{j1}, \cdots, w_{j,m-1})$.

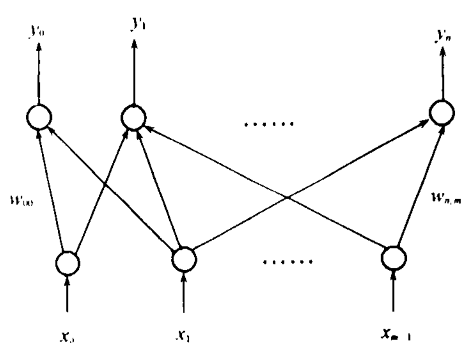

Figure 3 The structure of neural networks for state recognition.

3.2 UR&SAL (unsupervised resonance & supervised adjust learning) algorithm [2, 4]

UR&SAL is a learning algorithm of neural networks. Its process can be described as follows.

Step 1 Determine the range of new-class, indicated as $[S_S, S_B]$. Set the bounds of distance parameter ρ, expressed as $[\rho_S, \rho_B]$. Normalize x_i $(i=0, \cdots, m-1)$ and record the total number of class as C.

Step 2 Set $K=1$ (K is the sequence number of class).

Step 3 Let $L=\sum_{j=1}^{K} S_j$ (S_j stands for the number of subclass clustered from class j and L is the sequence number of new-class). Set $S_k=1$ and $W'_L=X(0)$, where $X(0)$ is the first input vector of class j.

Step 4 Execute the following steps to every input learning sample X.

(1) Activity. Compute the distance between X and W'_i with equation (2).

$$d'_i=\|X-W'_i\|=\Big[\sum_{j=0}^{m-1}(x_j-w_{ij})^2\Big]^{1/2} \tag{2}$$

(2) Competition. Find vector W'_i which corresponding to $d'_i=\min_i d'_i$. If there are many vectors, select one at random.

(3) Resonance. If $d'_i>\rho$, then adding a new node to the second layer and give it a new sequence number L. At the same time, set $W'_l=X$, $S_k=S_k+1$, and $L=L+1$ turn to step 5.

(4) Learning. If $d'_i\le\rho$, then set $W'_i=W'_i+\beta(X-W'_i)$, where β is the learning efficiency.

Step 5 If the learning samples of class K have been completely inputted, turn to step 6. Otherwise turn to step 4.

Step 6 If $S_K\in[S_S, S_B]$ is satisfied, turn to next step. Otherwise update ρ and turn to step 3.

Step 7 If $K\le C$, then set $K=K+1$ and turn to step 3. Otherwise turn to next step.

Step 8 Initialize the weight vector with equation $W_L(0)=W'_L$ $(L=1,2,\cdots,n)$, where n $(n=\sum_{i=1}^{C} S_i)$ stands for the total number of the new-class.

Step 9 Select a vector $X(t)$ from the set of input vector randomly and compute the distance d'_L between $X(t)$ and $W_L(t)$. If $d'_r=\min_L d(X(t), W_L(t))$, then $X(t)$ should be put in the new-class r. By virtue of the relationship between the new-class and class, we can infer that $X(t)$ belongs to class K'. In fact, $X(t)$ belonged to class K for it was a learning sample we had known.

Step 10 Update the weight vector $W_r'(t)$ with following equation:

$$\begin{cases} W_r'(t+1)=W_r'(t)+\alpha(t)[X(t)-W_r'(t)] & K=K' \\ W_r'(t+1)=W_r'(t)-\alpha(t)[X(t)-W_r'(t)] & K\neq K' \\ W_L'(t+1)=W_L'(t) & L\neq r \end{cases} \quad (3)$$

where $\alpha(t)$ is a step function such as $\alpha(t)=\alpha_0(1-t/T)$, where t is the learning steps, α_0 and T are constants and can be determined by user.

Step 11 If the learning samples have been completely inputted, turn to next step. Otherwise turn to step 9.

Step 12 Calculate the error with following equation:

$$E=\max_L[0.5\times\|W_L'(t+1)-W_L'(t)\|^2] \quad (4)$$

If the condition "$E<\varepsilon$" is satisfied, stop the recurrent process, otherwise turn to step 9, where ε is the error bounds which given by user.

From the above process, two conclusions can be drawn:

(1) Step 1–step 7 is the first stage of the learning, which is called unsupervised resonance. It is similar to ART (adaptive resonance theory) algorithm and guarantees that the new-class satisfies the presupposition.

(2) Step 8–step 12 is the second stage of the learning, which is called supervised adjust learning. Its main purpose is to force the weight vector to approach the central vector of the new-class. At the same time it can also rectify the error of the first stage.

4 Applications of the Improved Minimum Distance Method

In order to validate the feasibility of the improved minimum distance method, the method is applied to the recognition of the three-section heating furnace. After feature extraction, the feature space is composed of four variables. They are the temperature of heating section Ⅰ, the temperature of heating section Ⅱ, the temperature of soaking section and the pressure of furnace. Consequently there are four input nodes in this neural networks. The data on-site is classified to 5 classes denoted as $M=\{M_1, M_2, M_3, M_4, M_5\}$={"Normal", "High Temperature", "Low Temperature", "High Pressure" and "Low Pressure"}. There are 120 samples measured on-site altogether. Among them, 80 samples are treated as learning samples and 40 samples are used as testing data. The data on-site is not satisfied the presupposition discussed above. So MDM can not be utilized directly.

Here the improved minimum distance method (IMDM) is used to this recognition process. Set $[S_S, S_B]=[1, 10]$, $[\rho_S, \rho_B]=[0.001, 0.005]$ and $\rho=0.002$ because all the vectors have been normalized. After the implementation of step 1–step 7, it can be seen that M_1, M_4 and M_5 are composed of only one subclass, however M_2 and M_3 are composed of seven subclasses. For example, $M_2=\{M_{21}, M_{22}, M_{23}, M_{24}, M_{25}, M_{26}, M_{27}\}$={"Only the temperature of soaking section is high", "Only the temperature of heating section Ⅰ is high", "Only the temperature of heating section Ⅱ is high", "Both the temperatures of soaking section and heating section Ⅰ are high", "Both the temperatures of soaking section and heating section Ⅱ are high", "Both the temperatures of heating section Ⅰ and heating section Ⅱ are high", "the temperatures of soaking section, heating section Ⅰ and heating section Ⅱ are all high"}, and M_3 is similar to M_2 except it corresponds to the "Low Temperature". The clustering result shows that the first stage of UR&SAL algorithm is absolutely reasonable. So there are 17 (=1+7+7+1+1) output nodes in the second layer. In succession, we can determine the central vector W_i' $(i=1,\cdots,17)$ by use of step 8–step 12 of UR&SAL algorithm. The relationship between new-class and class in state recognition of heating furnace is listed in **table 2**.

Table 2 Relationship between new-class and class in state recognition of heating furnace

new-class	1	2–8	15	16	17
class	1	2	3	4	5
result	N (Normal)	HT(High temperature)	LT(Low temperature)	HP(High pressure)	LP(Low pressure)

Then the neural networks can be utilized to state recognition of heating furnace. 40 testing samples and their recognition result are listed in **Appendix 1**.

5 Conclusions

From Appendix 1, it can be seen that the accuracy of recognition is up to 97.5%. When MDM is applied to recognition, the accuracy of recognition is only 35%. So we can say that the recognition accuracy of IMDM is higher enough than MDM. The recognition rate of IMDM is also faster than MDM because the former is executed by neural networks. In a word, the improved minimum distance method is very effective.

References

[1] Z. Q. Bian. *Pattern Recognition* (in Chinese) [M]. Tsinghua University Press, Beijing, 1988.

[2] Q. Li. *State Recognition and Establishment of the Fast Inference Theory without Searching* (in Chinese) [D]. Beijing: Univ. of Sci. & Tech. Beijing, 1995.

[3] H. B. Shi. *Neural Networks and Its Applications* (in Chinese) [M]. Xi an Transportation University Press, Xi an, 1993.

[4] S. J. Wang, S. L. Wang. *Computer Research and Development* (in Chinese) [J]. 32 (1995), No. 4, p. 6.

Appendix 1 The recognition result of 40 testing samples

No.	Soaking/℃	Heating Ⅰ/℃	Heating Ⅱ/℃	Pressure/Pa	New-class	Class	Result	Correctness
1	1 250	1 300	1 250	19.6	1	1	N	Δ
2	1 243	1 200	1 157	19.6	15	3	LT	Δ
3	1 281	1 325	1 263	22.6	1	1	N	Δ
4	1 309	1 354	1 286	24.5	1	1	N	Δ
5	1 283	1 322	1 266	16.7	17	5	LP	Δ
6	1 267	1 315	1 266	23.5	1	1	N	Δ
7	1 352	1 396	1 357	21.6	8	2	HT	Δ
8	1 246	1 294	1 258	23.5	1	1	N	Δ
9	1 272	1 335	1 269	22.6	1	1	N	Δ
10	1 306	1 355	1 283	27.5	16	4	HP	Δ
11	1 251	1 305	1 259	22.6	1	1	N	Δ
12	1 248	1 314	1 265	21.6	1	1	N	Δ
13	1 165	1 310	1 260	24.5	9	3	LT	Δ
14	1 286	1 320	1 260	29.4	16	4	HP	Δ
15	1 290	1 311	1 273	23.5	1	1	N	Δ
16	1 278	1 327	1 265	22.6	1	1	N	Δ
17	1 283	1 330	1 284	18.6	17	5	LP	Δ
18	1 258	1 306	1 263	21.6	1	1	N	Δ
19	1 275	1 315	1 255	22.6	1	1	N	Δ
20	1 245	1 320	1 270	21.6	1	1	N	Δ
21	1 320	1 336	1 304	22.6	6	2	HT	Δ
22	1 258	1 300	1 254	21.6	1	1	N	Δ
23	1 270	1 285	1 239	24.5	1	1	N	Δ
24	1 257	1 287	1 274	20.6	1	1	N	▲
25	1 256	1 322	1 258	23.5	1	1	N	Δ
26	1 275	1 325	1 265	21.6	1	1	N	Δ
27	1 220	1 280	1 270	21.6	12	3	LT	Δ
28	1 265	1 330	1 265	20.6	1	1	N	Δ
29	1 270	1 329	1 278	21.6	1	1	N	Δ
30	1 250	1 300	1 250	17.7	17	5	LP	Δ
31	1 259	1 299	1 255	22.6	1	1	N	Δ
32	1 258	1 308	1 265	21.6	1	1	N	Δ
33	1 239	1 325	1 180	23.5	13	3	LT	Δ
34	1 255	1 305	1 255	20.6	1	1	N	Δ
35	1 279	1 328	1 245	23.5	1	1	N	Δ
36	1 264	1 309	1 255	21.6	1	1	N	Δ
37	1 270	1 335	1 276	14.7	17	5	LP	Δ
38	1 248	1 315	1 275	22.6	1	1	N	Δ
39	1 250	1 306	1 254	22.6	1	1	N	Δ
40	1 243	1 320	1 267	20.6	1	1	N	Δ

Note: Δ indicates that the recognition result is right and ▲ indicates that the recognition result is wrong.