http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 11 (8): 1091-1096, 2012
ISSN 1812-5638 / DOI: 10.3923/itj.2012.1091.1096

Classifier Design Algorithms Aimed at Overlapping Characteristics

Li Yi-Bing, Li Jing-Chao and Kang Jian
Institute of Communication Engineering, Harbin Engineering University, Harbin 150001, China

Abstract: In the complex electromagnetic environment, the feature parameters could be classified with favorable separating degree in high Signal-to-noise ratio (SNR) by conventional gray relation algorithm. In low SNR, however, it will be existed overlapping phenomenon and even hard to select standard sample values of unstable parameter. Aiming to these issues, this study does further research on the classifier design method based on gray relation algorithm. By introducing the mean samples gray relation, the selection of standard sample value is improved. Moreover, the features are weighed according to the surplus degree which increases the adaptive ability of the gray relation algorithm. Meanwhile, the paper used the interval relation classifier based on adaptive entropy weight to classify the features with overlapping properties. It overcome the problem of selecting standard samples and had some adaptive ability to some extent. Comparing with other algorithms, it has the best classification effect.

Key words: Classifier design algorithms, overlapping characteristics, gray relation algorithm, interval relation classifier

INTRODUCTION

Nowadays, more and more signals and noise are exiting and the communication environment is more and more complex. In order to get high recognition rate, the extraction of relatively stable characteristic parameters are important under the environments of low SNR and the choice of classifier design is also very critical. The adaptive ability of general traditional classifier (Lee *et al*., 2010a; Foo *et al*., 2011; Nourain and Samir, 2011) is poor to some extent. In the low SNR environments, even if the characteristic parameters are stable, it is also hard to achieve a satisfactory result. The support vector machine (Wu and Zhao, 2006; Shafri *et al*., 2009; Yeary *et al*., 2009; Lee *et al*., 2010b; Huang *et al*., 2010) also has better recognition effect for signal feature classification. However, the selection of kernel function is the key process of recognition, it directly affects the final result. The proposed of neural network classifier (Qureshi and Jalil, 2002; Tahir *et al*., 2006; Lotfi *et al*., 2010; Razavi and Tolson, 2011) open up a very good road for the design of classifiers. It is more widely used in the signal identification field because of its good adaptive ability. The literature (Zhang *et al*., 2005) recognized several kinds of radar signals by using neural network classifier based on the characteristic parameters of similar coefficient. It achieved a good recognition effect to the parameters which had good gathered degrees but for the signals which had poor gathered degrees, the recognition rate was falling rapidly. The gray relation theory is also more and more widely used in the classifier design in recent years. Yun and Jing-Chao (2011) put forward a radar signal recognition algorithm based on neural network and gray relation theory and it also put up some theoretical research of blind recognition algorithm based on neural network and gray relation theory. However, the algorithm only suitable for the signals which have good gathered characteristics and it is not applicable for the parameters with overlapping features. The interval relation algorithm (Wei, 2006) is always used in the decision problem analysis (Zhang, 2005) and target selection (Chen *et al*., 2010). While, the selection of relation coefficient weights are the key problem of interval relation algorithm identification.

Based on the several characteristics of classifiers proposed above, this study analyzes 5 classifier design algorithm aimed at the signals with overlapping characteristics parameters. It focuses on the study of gray relation algorithm and proposed some corresponding improved algorithms. It also compares and simulates the classification effect of the overlapping characteristics parameters using the above algorithm in different SNR. It provides a good theoretical basis for the recognition of overlapping characteristic parameters in engineering application.

NEURAL NETWORK CLASSIFIER

In recent years, the neural network technology develops rapidly. It is a complex processing unit which is

Corresponding Author: Li Jing-Chao, Institute of Communication Engineering, Harbin Engineering University, Harbin 150001, China

widely connected by a number of simple, processing units. It is also a highly complex nonlinear dynamic system reflecting many basic characteristics of human brain function.

The neural network has powerful pattern recognition ability, good automatically adaptive ability to the changes of environment and it is able to handle complex non-linear identification problem well. With its strong robustness and potential fault tolerance, the neural network has been more widely used in the design of signal classifier.

Kasakawa *et al*. (2010) made a more detailed description of the basic theory of the neural network, so this study is no longer retelling. The paper mainly compares several improved gray relation classifiers and the neural network classifier with strong adaptive ability. It puts the emphasis on the further discussion of gray relation classifiers. The recognition effects of different classifier designs are explored in low SNR when the feature parameters have the overlapping performance. Meanwhile, it discusses and chooses the classifier with best qualification in the light of different situation.

GRAY RELATION CLASSIFIER

The ordinary gray relation algorithm: People often use the depth of color to describe the clear degree of the information. Usually, the “black” means the information is unknown, the “White” means the information is completely clear. And the “gray” means part of the information is clear and part of the information is not clear. So, the incomplete information system is often called “the gray system”.

The gray relational analysis is a quantitative description and comparative method to the change trend of the system development. Suppose the system behavior sequence is:

$$X_0 = (x_0(1), x_0(2), \ldots, x_0(n))$$
$$X_1 = (x_1(1), x_1(2), \ldots, x_1(n))$$
$$X_i = (x_i(1), x_i(2), \ldots, x_i(n))$$
$$X_m = (x_m(1), x_m(2), \ldots, x_m(n))$$

For $\xi \in (0, 1)$:

$$\gamma(x_0(k), x_i(k)) = \frac{\min_i \min_k |x_0(k) - x_i(k)| + \xi \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \xi \max_i \max_k |x_0(k) - x_i(k)|} \quad (1)$$

$$\gamma(X_0, X_i) = \frac{1}{n}\sum_{k=1}^{n} \gamma(x_0(k), x_i(k)) \quad (2)$$

where, ξ is the distinguish coefficient and the value is usually 0.5. $\gamma(X_0, X_i)$ is the gray relation degree of X_0 and X_i. The gray relation degree $\gamma(X_0, X_i)$ is usually brief written down for γ_{0i}, the relation coefficient of the k points $\gamma(x_0(k), x_i(k))$ is usually brief written down for $\gamma_{0i}(k)$.

The improved mean value samples gray relation algorithm: It can directly use ordinary gray relation theory to calculate the gray relation degree for the signals with the characteristics of a standard sample values. However, we often can’t get certain standard parameters of the signals in the complex electromagnetic environment. The parameter values of the signals are often fluctuant in a certain range. In this case, the selection of the standard sample values becomes the key problem of the classifier design. Usually, we will take a possible characteristic value as the reference value for the recognition and classification of the features. However, on this occasion, the selection of the reference characteristic value has certain randomness. The selection of characteristic value has a direct influence on the final classifier recognition results. Thus, the study introduces the mean samples gray relation algorithm.

Suppose the characteristic value x is an element of the characteristic parameters set X of a signal. The possible characteristic values are $\{x_i\} = \{x_1, x_2, \ldots, x_n\}$, while, $x_1 \neq x_2 \neq \ldots \neq x_n$, thus, take the mean value of all the n possible characteristic values as the standard reference characteristics, that is:

$$x_0 = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

From the analysis of the theory, compared with the method which randomly selected the reference samples, choosing the mean characteristic values as the reference sequence has better stability. Of course, when the overlaps of different signal parameter values are too big, that is the mean values of the samples is adjacent, the recognition rate of the algorithm is dropped rapidly. Therefore, this algorithm has certain requirements for the distance of the samples.

The improved adaptive mean value samples gray relation algorithm: Yun *et al*. (2010) put forward gray relation algorithm based on adaptive entropy weight. For the calculation of the features’ relation degree, it introduces the information theory into the calculation of the relation degree formula. Based on the improved mean value samples gray relation algorithm, the improved adaptive mean value samples gray relation algorithm combine the concepts of surplus degree in the information theory of literatures (Yun *et al*., 2010). First process the characteristic distance Δx_{ij} as follows:

$$p_{ij} = \Delta x_{ij} \Big/ \sum_{i=1}^{M} \Delta x_{ij} \tag{3}$$

where, M express the categories of the signals. Define the entropy value:

$$E_j = -\sum_{i=1}^{M} p_{ij} \ln p_{ij}$$

and the maximum entropy value: E_{max} = InM, the relative entropy value: ej = E_j/E_{max}.

Reference the concepts of surplus degree in information theory, the definition of the surplus degree of the j element is:

$$D_j = 1 - e_j \tag{4}$$

The surplus degree actually means removing the difference of the j entropy characteristics from the optimal entropy characteristics. The bigger of D_j, the more importance of the characteristics and it should be given greater weight.

Finally calculates the weight $a_i(j)$ of the characteristics j:

$$a_i(j) = D_j \Big/ \sum_{j=1}^{N} D_j \tag{5}$$

While:

$$\sum_{j=1}^{N} a_i(j) = 1,\ a_i(j) \geq 0$$

Let the weight coefficients multiply with the corresponding relation coefficients, it will get the relation degree value.

From the perspective of statistics, the characteristics with big deviation will be the best to react the difference among different types, so, the bigger the difference, the more important the characteristic. Give the different characteristics for the different weights. For the characteristics with large difference, it will get large weight. Accordingly, the characteristics with small difference will get small weight. So, it increases the adaptive ability of the gray relation algorithm to some extent, it is beneficial to final classification and identification.

The interval gray relation algorithm based on entropy weight: The characteristic value extracted is often not a certain value due to the complex communication environment but always change in a certain range. Thus, the paper introduces the concept of interval gray relation. The process of the algorithm is described as follows:

First define the features interval matrix as:

$$S = \begin{pmatrix} [s_{11}^{min}\ s_{11}^{max}] & [s_{12}^{min}\ s_{12}^{max}] & \cdots & [s_{1n}^{min}\ s_{1n}^{max}] \\ [s_{21}^{min}\ s_{21}^{max}] & [s_{22}^{min}\ s_{22}^{max}] & \cdots & [s_{2n}^{min}\ s_{2n}^{max}] \\ \vdots & \vdots & \vdots & \vdots \\ [s_{m1}^{min}\ s_{m1}^{max}] & [s_{m2}^{min}\ s_{m2}^{max}] & \cdots & [s_{mn}^{min}\ s_{mn}^{max}] \end{pmatrix} \tag{6}$$

While, m expresses the category of the signals, n expresses the number of the characteristic parameters. Suppose the characteristics interval n of a signal to be identified is $[s_{0n}^{min}\ s_{0n}^{max}]$, so, the interval leave degree of it with the known signal features interval can be defined as:

$$d_{mn} = \frac{1}{\sqrt{2}}\sqrt{(s_{mn\,min} - s_{0n\,min})^2 + (s_{mn\,max} - s_{0n\,max})^2} \tag{7}$$

According to the gray relation theory, the interval relation coefficient is expressed as:

$$\xi_{mn} = \frac{\min_m \min_n \{d_{mn}\} + \rho \max_m \max_n \{d_{mn}\}}{d_{mn} + \rho \max_m \max_n \{d_{mn}\}} \tag{8}$$

This calculation results can constitute an interval relation coefficient matrix:

$$\xi_{mn} = \begin{pmatrix} \xi_{11} & \xi_{12} & \cdots & \xi_{1n} \\ \xi_{21} & \xi_{22} & \cdots & \xi_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \xi_{m1} & \xi_{m2} & \cdots & \xi_{mn} \end{pmatrix} \tag{9}$$

The same as the above process, according to the concept of surplus degree in 2.3, get the entropy weight of each relation coefficient and multiply them with the corresponding relation coefficient, so, the relation degree value can be gotten finally.

It can be known from the theoretical analysis, the interval relation algorithm based on entropy weight has many advantages. On the one hand, it has certain adaptive ability to adjust the characteristic weight coefficient and on the other hand, it can classify and identify the characteristic parameters with overlapping properties in low SNR. So it has better resistance ability for the overlapping of feature interval.

THE SIMULATION RESULTS AND ANALYSIS

The study mainly analyses five classifier design method, while, it put the emphasis on the classifier design algorithm based on the gray relation theory. It classifies and identifies the characteristic parameters with overlapping properties and compares these methods with the neural network classifier with strong adaptive ability, then calculates the recognition rate and the computation time of different classifications.

First, the study uses the two characteristic parameters of 6 signals to constitute a 2d characteristic vector. They have good separation degrees in high SNR environments but have certain overlapping properties under low SNR. The distributions of the characteristic parameters are shown in Fig. 1 and 2. While, Fig. 1 shows the 2d features distribution of 6 kinds of signals under the SNR of 20 db and Fig. 2 shows the 2d features distribution of 6 kinds of signals under the SNR of 2 db.

Using the theoretical analysis about the 5 classifier design method analysed before, classify and identify the characteristics of the six signals under different SNR environments and the recognition rate results are shown in Table 1.

It can be seen from Fig. 1, when the SNR is 20 db, the 5 kinds of classifiers all have very good recognition properties. Combined with Fig. 1 to analyze the results, when the characteristic parameters of the 6 signals have good separation degree, the classifiers can all achieve the recognition rate of 100%. When the SNR is gradually reduced, the recognition rates of the ordinary gray relation algorithm, the improved mean value samples gray relation algorithm and the improved adaptive mean value samples gray relation algorithm, are all gradually declining with the SNR reduced. Because the strong adaptive ability of the neural network classifier, it is not so sensitive to the reducing of SNR. But when the SNR is 2 db, the recognition rate is reducing to less than 90%. But for the interval gray relation algorithm based on entropy weight, even if the SNR is 0 db, it still maintain a high recognition rate, Combined with Fig. 2 to analysis the results, When there are some overlapping part of the signal characteristics parameters, especially when overlap is more serious, the interval gray relation algorithm based on entropy weight has the best recognition effects.

In order to measure the complex properties of different algorithms, the simulation time of different classifiers is shown in Table 2.

It can be seen from the comparison in the table that, the calculation time of the classifier designs based on gray relation algorithm have little difference, that is to say, these classifiers are equal in the complexity of algorithm. For the neural network classifier, the simulation time is long to some extent, it is decided by the process of training to realize the signal recognition. However, usually, only one time training is needed, the process of training is eliminating in the following recognition procedure. From this perspective, the difference of complexity degrees among these algorithms will reduce relatively compared with the other several gray relation algorithms.

Neural network classifier has been more widely used for its stronger adaptive ability. However, for the

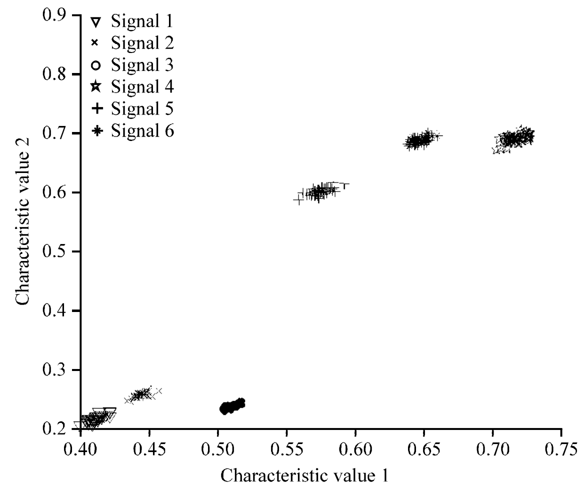

Fig. 1: The 2 D features distribution of 6 kinds of signals under the SNR of 20 db

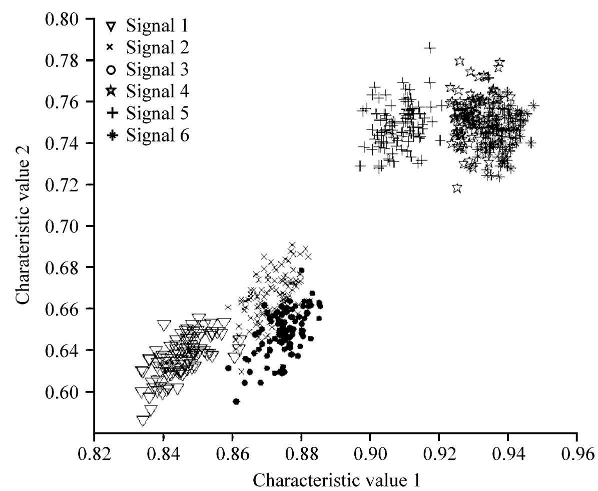

Fig. 2: The 2 D features distribution of 6 kinds of signals under the SNR of 2 db

Table 1: The recognition rate of 5 classifier design methods in different SNR

	20 db	10 db	5 db	2 db	0 db
Neural network classifier (%)	100	100	96.7	87.3	71.3
Ordinary gray relation algorithm (%)	100	98	91.3	83.5	67.8
The improved mean value samples gray relation algorithm (%)	100	99.3	88.7	75.7	67.8
The improved adaptive mean value samples gray relation algorithm (%)	100	99.5	95.7	88.5	76.2
The interval gray relation algorithm based on entropy weight (%)	100	100	100	95.33	92.33

Table 2: The simulation time of different classifier design algorithms

Algorithm name	Simulation time (sec)
Neural network	4.97
Gray relation algorithm	1.03
The improved mean value samples gray relation algorithm	1.02
The improved adaptive mean value samples gray relation algorithm	1.02
The interval gray relation algorithm based on entropy weight	1.02

characteristic parameters with overlapping properties, using neural network classifier to classify is not the optimal choice. From the compared results of the recognition rate and the simulation time in the Table 2, it will have better recognition result using interval gray relation algorithm based on entropy weight for the characteristic parameters with overlapping properties.

CONCLUSION

In order to solve the classification problem of the signals with overlapping characteristic parameters, this paper puts forward five classifier design methods. And the paper mainly did further research on the gray relation theory, proposed three kinds of improved gray relation theory algorithm, discussed the classification performance of different algorithms aimed at the identification of characteristic parameters with overlapping properties and simulated the experiments and compared them with the neural network classifier which has stronger adaptive ability. It can be seen from the theoretical analysis and the simulation results that, although the neural network classifier has stronger adaptive ability, when the characteristic parameters of signals have overlapping properties in low SNR, the interval gray relation algorithm based on entropy weight has the best classification effect. This provides a good theoretical basis for the classification of signals with overlapping characteristics in engineering practice.

ACKNOWLEDGMENTS

This study is sponsored by the National Basic Research Program of China No. 61393010101-1 and Fundamental Research Funds for the Central Universities No. HEUCF100810.

REFERENCES

Chen, G., X. Liu and S. Guo, 2010. Selection method of radar countermeasures targets based on interval number grey relation analysis. Electron. Inform. Warfare Technol., 25: 57-60, 68.

Foo, B., D.S. Turaga, O. Verscheure, M. van der Schaar and L. Amini, 2011. Configuring trees of classifiers in distributed multimedia stream mining systems. IEEE Trans. Circuits Syst. Video Technol., 21: 245-258.

Huang, C.C., R.G. Chung, R.C. Chen, T.S. Chen, T.N. Le, C.J. Hsu and Y.C. Tsai, 2010. Finding an optimal combination of key training items using genetic algorithms and support vector machines. Inform. Technol. J., 9: 652-658.

Kasakawa, T., H. Tabata, R. Onodera, H. Kojima, M. Kimura, H. Hara and S. Inoue, 2010. An artificial neural network at device level using simplified architecture and thin-film transistors. IEEE Trans. Electron Devices, 57: 2744-2750.

Lee, L.H., C.H. Wan, T.F. Yong and H.M. Kok, 2010a. A review of nearest neighbor-support vector machines hybrid classification models. J. Applied Sci., 10: 1841-1858.

Lee, L.H., D. Isa, W.O. Choo and W.Y. Chue, 2010b. Tournament structure ranking techniques for Bayesian text classification with highly similar categories. J. Applied Sci., 10: 1243-1254.

Lotfi, A., K. Mezzoug and A. Benyettou, 2010. Rotated kernel neural networks for radar target detection in background noise. J. Applied Sci., 10: 1331-1335.

Nourain, N. and B.B. Samir, 2011. Statistical model and wavelet function for face recognition. J. Applied Sci., 11: 1213-1218.

Qureshi, I.M. and A. Jalil, 2002. Object recognition using ANN with backpropogation algorithm. J. Applied Sci., 2: 281-287.

Razavi, S. and B.A. Tolson, 2011. A new formulation for feed forward neural networks. IEEE Trans. Neural Networks, 22: 1588-1598.

Shafri, H.Z.M. and F.S.H. Ramle, 2009. A comparison of support vector machine and decision tree classifications using satellite data of Langkawi Island. Inform. Technol. J., 8: 64-70.

Tahir, N.M., A. Hussain, S.A. Samad, H. Husain and M.M. Mustafa, 2006. Eigenposture for classification. J. Applied Sci., 6: 419-424.

Wei, G., 2006. New method of gray relational analysis to multiple attribute decision making with intervals. Syst. Eng. Electron., 28: 1358-1359,1383.

Wu, F. and Y. Zhao, 2006. Least squares support vector machine on Morlet wavelet Kernel function and its application to nonlinear system identification. Inform. Technol. J., 5: 439-444.

Yeary, M.B., S. Nemati, T.Y. Yu, Y. Wang and Y. Zhai, 2009. A support-vector-machine-based approach to rf sensor spectral signature classifications. IEEE Trans. Instrument. Measure., 58: 221-228.

Yun, L. and L. Jing-Chao, 2011. Radar signal recognition algorithms based on neural network and grey relation theory. Proceedings of the Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC), July 26-30, 2011, Harbin, pp: 1482-1485.

Yun, L., S. Xi-Cai and Z. Ruo-lin, 2010. Application of improved gray relation algorithm on radiation source recognition. J. Commun., 31: 166-171.

Zhang, G.X., J. Wei-Dong and H. Lai-Zhao, 2005. Resemblance coefficient based feature selection algorithm for radar emitter signal recognition. Signal Process., 21: 663-667.

Zhang, J.J., 2005. The method of grey related analysis to multiple attribute decision making problems with interval numbers. Syst. Eng. Electron., 27: 1030-1033.