Research on Recognition Methods of underwater acoustic signal based on higher-order statistics

Zhenyu Li[a], Zhiqiang Chen[b], Bo Zhang[c], Zhihui Zhang[d] and **Ziyang Yu[e]**

Dalian Scientific Test and Control Technology Institute, Dalian 116013, China

[a]Lizhenyu@163.com, [b]Chenzhiqiang@163.com, [c]Zhangbo123@163.com, [d]Zhangzhihui@163.com, [e]34387344@qq.com

Abstract: In acoustic signal recognition problem, different analysis methods can extract different characteristics for the same goal. Good feature fusion methods can take advantage of different traits of each feature, complement each other ,remove redundancy, get more robust new features, and improve algorithm recognition rate. This method can also complete the data compression dimensionality reduction and improve the real-time algorithm. This shows that feature fusion is extraordinary. This paper presents a signal recognition method based on high order statistics and power spectrum estimation and theoretical simulation results. This method can be used to recognize underwater acoustic signals and has a high recognition rate.

Keywords: Higher order statistics; Bispectrum; Recognition; Typical correlation analysis method..

1. Introduction

Hydroacoustic target recognition belongs to the category of pattern recognition. This problem has been widely concerned in the field of underwater acoustic engineering[1].As shown in Figure 1 below,classical object recognition has its own mature theoretical framework[2] .
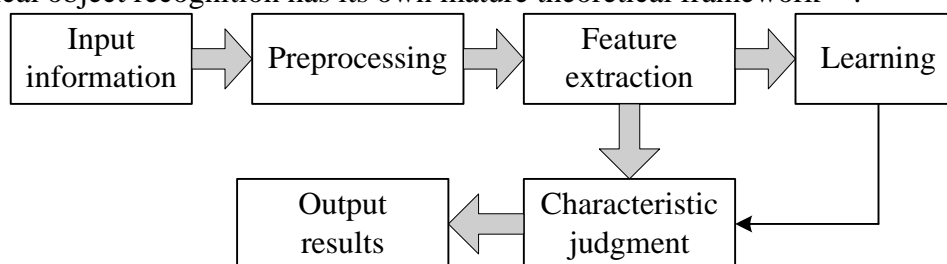

Fig. 1 pattern recognition framework

In the pattern recognition problem, the quality of the feature extraction results directly determines the effectiveness of the target recognition. At present, a large number of research work focuses on how to extract the most suitable features for classification. Power spectrum feature is the most basic feature of the signal.And it is a frequency domain analysis. There are also time-frequency-domain-based time spectra such as wavelet transforms and Hilbert-Huang transforms. The high-order spectrum of signal accumulation can be used to extract the characteristics of the target signal. These spectral methods have also become a common means of studying the characteristics of underwater acoustic signals.

There are many features to choose from, including richness. Too few features can not reflect the goal.Too many features introduce both noise interference data and over-learning. In order to ensure the irrelevance of the features, the related features should be de-correlated and reflect the target information as much as possible. Therefore, this paper proposes an underwater acoustic recognition method based on higher-order statistics and power spectrum estimation.

In the case of submarine classification, different subjects and different starting points for research lead to differences in the choice of feature set. Some characterization studies have matured and cured. Some new features are still being constantly explored. In this paper, in order to meet the requirements of algorithm verification, the feature set meets the following three points at least:

(1)The wide is a feature that reflects as much as possible aspects of the target's characteristics.This can play the advantages of feature fusion.

(2)Accurate is the extracted features that should be as accurate as possible to reflect a certain characteristic of the target and possess the maximum discernibility.There are two layers of meaning.One is extracting accurately in the selection of the original features.The second is the processing of primitive features. For example, in the process of feature fusion and optimization, the resulting comprehensive features are as accurate as possible.

(3)Less is the use of the smallest amount of data to express as much feature information as possible.At the same time, less also means that the correlation between features is the smallest.

The contradiction between wide, accurate, and less is the key difficulty in target recognition.This article emphasizes how to use information fusion and feature optimization methods to achieve the balance of the three.As described in 2.3.2, not all features are suitable for using the typical correlation analysis method algorithm.In other words, the effects of different feature selection algorithms are different.In order to study the applicability of different features to this algorithm, different characteristics of underwater acoustic signals need to be extracted.

In general, the measured signal of underwater acoustic engineering is usually non-linear non-Gaussian non-stationary.Traditional signal processing emphasizes smoothness, randomness, and linearity. For complex signal processing, there is often a reasonable approximation to simplify the problem.Although sometimes it brings some convenience, these approximate methods sometimes have serious flaws.For example, some signals tend to lose some features after linear approximation and reduce the accuracy of signal description.This will affect the accuracy and reliability of target recognition.For the processing of non-Gaussian signals, an analysis method of high-order statistics, which has recently been frequently mentioned, is becoming a powerful tool for analyzing and processing non-Gaussian nonlinear signals;In addition, time-frequency analysis methods such as the short-time Fourier transform, wavelet transform, and Wigner distribution have achieved better results than traditional Fourier transforms in processing non-stationary signals.

In this paper, what needs to be analyzed is the effect on the algorithm when different feature combinations are combined.Spectral analysis, as the most basic feature of the signal, is of course taken into account when analyzing the effect of the feature on the algorithm.It is obvious and necessary to use the most basic power spectrum as the feature to be extracted and observe its effect on the algorithm.Based on the above considerations, the three major features extracted in this paper are power spectrum features, bispectrum features, and Wigner distribution features.Using the fewest number of features achieve the purpose of screening features from multiple levels and multiple angles and observing the effect of different feature combinations on the algorithm as much as possible.It is widely known that the power spectrum feature is the most basic and simple signal feature.so its theory is not analyzed here.This article focuses on bispectral features and Wigner distribution features.

2. Theoretical analysis

2.1 Bispectral features

Higher-order statistics are statistics with orders greater than second-order.Higher-order statistics mainly include high-order moments, higher-order cumulants, and high-order cumulant spectra (abbreviated as high-order spectra).Because higher-order statistics contain a wealth of information that second-order statistics do not have.Any problem that has been analyzed and processed by the power spectrum and related functions without satisfactory results is worthy of re-examination of higher-order statistical methods.Higher-order statistics are based on second-order statistics.The power spectrum and spectrum mentioned here are all second-order statistics.Second-order statistics

usually do not contain phase information.But higher-order statistics overcome this shortcoming and contain richer information[4].

It can be seen that the high-order cumulant of any Gaussian process is zero.But higher moments are not the same.The third-order moment of a zero-mean Gaussian random variable is zero, but non-Gaussian random variables are not necessarily the case. In theory, using high-order cumulants can completely suppress the effects of Gaussian colored noise.

Among the higher-order cumulants, the most commonly used is the third-order cumulant known as trispectrum or bispectrum.In the higher order spectrum, the lowest order bispectrum is the easiest to process and contains all the features of the higher-order spectrum.Bispectrum can extract the phase information of the signal.Bispectrum represents random signals from higher order probabilistic structures and compensates for the shortcomings of second-order statistics in this area.

Higher spectrum is defined by cumulants.Assuming a zero-mean order stationary process, the order cumulant is defined as follows:

$$C_{kx}(\tau_1, \tau_2, \cdots, \tau_{k-1}) = sum\{x(n), x(n + \tau_1), \cdots, x(n + \tau_{k-1})\} \tag{1}$$

We suppose that high-order cumulants $C_{kx}(\tau_1, \tau_2, ..., \tau_{k-1})$ are absolutely corollable:

$$\sum_{\tau_1 = -\infty}^{+\infty} \cdots \sum_{\tau_k = -\infty}^{+\infty} \left| C_{kx}(\tau_1, \tau_2, \cdots, \tau_{k-1}) \right| < \infty \tag{2}$$

We define the k-order spectrum as k-dimensional Fourier transform of k-order cumulant:

$$S_{kx}(\omega_1, \omega_2, \cdots, \omega_{k-1}) = \sum_{\tau_1 = -\infty}^{+\infty} \cdots \sum_{\tau_k = -\infty}^{+\infty} C_{kx}(\tau_1, \tau_2, \cdots, \tau_{k-1}) \exp\left[-j\sum_{i=1}^{k-1} \omega_i \tau_i\right] \tag{3}$$

In the special case of common power spectrum (2nd-order cumulant at k=2):

$$P(\omega) = S_{2x}(\omega_1) = \sum_{\tau_1 = -\infty}^{+\infty} C_{2x}(\tau_1) \exp\{-i(\omega_1 \tau_1)\} \tag{4}$$

Bispectrum is a special case when k=3

$$B_x(\omega_1, \omega_2) = S_{3x}(\omega_1, \omega_2) = \sum_{\tau_1 = -\infty}^{+\infty} \sum_{\tau_2 = -\infty}^{+\infty} C_{3x}(\tau_1, \tau_2) \exp\left[-i(\omega_1 \tau_1 + \omega_2 \tau_2)\right] \tag{5}$$

2.2 Canonical Correlation Analysis

The typical correlation analysis method studies the correlation coefficient between two vectors[5].The first problem solved by the canonical correlation analysis method can be simply described as:for the input vector $x \in \Re^p$ and $y \in \Re^q$,get two projection vectors $W_x \in \Re^p$ and $W_y \in \Re^p$ .Maximize the correlation coefficient ρ between a pair of projected variables $w_x^T x$ and $w_y^T y$ .As shown in Figure 2:
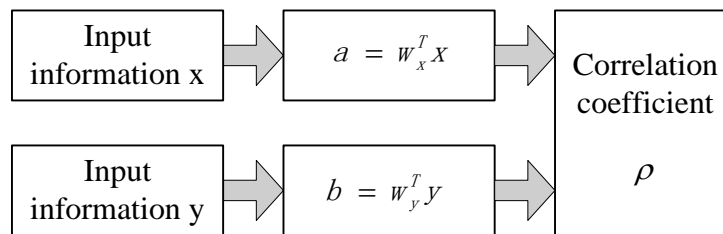

Fig 2 Typical Correlation Analysis Basic Process

The following is a detailed discussion of the algorithm[6-7].For observation sample sets $\{(x_i, y_i)\}_{i=1}^n \in \Re^P \times \Re^q$, and $\{x_i\}_{i=1}^n$ are $\{y_i\}_{i=1}^n$ two vector sets to be analyzed.They are recorded as $X = [x_1, ..., x_n] \in \Re^{p \times n}$ and $Y = [y_1, ..., y_n] \in \Re^{q \times n}$ .We believe that any pair of samples (x, y) comes from the collection $\{(x_i, y_i)\}_{i=1}^n$ .To simplify the problem, we assume that the sample $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i = 0$ and $\bar{y} = \frac{1}{n}\sum_{i=1}^n y_i = 0$ has been centralized.Let C_{xx} and C_{yy} be the covariance matrixes

of sample spaces X and Y[55] (or the overall scatter matrix). C_{xy} represents the cross-covariance matrices between them. $C = \begin{bmatrix} Var(x) & Cov(x,y) \\ Cov(y,x) & Var(y) \end{bmatrix} = \begin{bmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{bmatrix}$;Where C_{xx} and C_{yy} are positive definite matrices($C_{xy}^T = C_{yx}, r = rank(C_{xy})$).

For the following vectors:

$x_i = [x_{i1}, x_{i2},...,x_{ip}]^T$, $y_i = [y_{i1}, y_{i2},...,y_{iq}]^T$, $w_x = [w_{x1}, w_{x2},...,w_{xp}]^T$, $w_y = [w_{y1}, w_{y2},...,w_{yq}]^T$,We consider their linear combination;

$a = w_x^T x_i = w_{x1}x_{i1} + w_{x2}x_{i2} +...+ w_{xp}x_{ip}$ and $b = w_y^T y_i = w_{y1}y_{i1} + w_{y2}y_{i2} +...+ w_{yq}y_{iq}$ .The first goal of the canonical correlation analysis method is the projection vector of the sample set $w_x \in \Re^p$ and $w_y \in \Re^q$ .Maximum correlation between random variables $a = w_x^T x_i$ and $b = w_y^T y_i$ .Then we can establish the expression of a and b correlation coefficients.

$$\rho(w_x, w_y) = \rho'(a,b) = \frac{E[ab]}{\sqrt{E[a^2]E[b^2]}} = \frac{E[w_x^T xy^T w_y]}{\sqrt{E[w_x^T xx^T w_x]E[w_y^T yy^T w_y]}}$$

$$= \frac{w_x^T C_{xy} w_y}{\sqrt{w_x^T C_{xx} w_x \cdot w_y^T C_{yy} w_y}}$$

(6)

Where $E[\cdot]$ is the mathematical expectation. $C_{xx} = E[xx^T] = XX^T \in \Re^{p \times p}$, $C_{yy} = E[yy^T] = YY^T \in \Re^{q \times q}$ represent a covariance matrix within the set. $C_{xy} = E[xy^T] = XY^T \in \Re^{p \times q}$ represents an inter-set covariance matrix($C_{yx} = E[yx^T] = C_{xy}^T$).We use the statistical average of observations as a reasonable approximation of its mathematical expectation.This reasonable approximation has no critical effect on the final result of the algorithm.To sum up, the first goal of the canonical correlation analysis method is to maximize formula (1).For formula 1, there are the following properties:

$$\rho(kw_x, lw_y) = \rho(w_x, w_y) \qquad \forall k,l \in \Re$$

(7)

The extremes of $\rho(w_x, w_y)$ have nothing to do with the size of w_x and w_y , but only with their direction.

Obviously, if w_x and w_y are not constrained, w_x and w_y have infinitely many when the correlation coefficient reaches its maximum value.Let $w_x^T C_{xx} w_x = w_y^T C_{yy} w_y = 1$ to make the solution uniquely determined.

Thus, the basic problems of the canonical correlation analysis method can be summarized as:

Under the condition of formula (7), we must know the value of w_x and w_y to make the correlation coefficient $\rho(w_x, w_y)$ reach the maximum (resolve the constrained optimization problem

$$\max_{w_x, w_y} \rho(w_x, w_y) = w_x^T C_{xy} w_y \quad s.t \quad w_x^T C_{xx} w_x = 1, w_y^T C_{yy} w_y = 1$$

(8)

The projection vector (w_{1x}, w_{1y}) is obtained from Formula (3) as the first pair of projection vectors.At the same time we can find a corresponding combination (ρ_1, a_1, b_1) .The second step of the typical correlation analysis needs to consider is:Finding the second projection vector (w_{2x}, w_{2y}) .

$$w_{1x}^T C_{xx} w_{2x} = w_{1y}^T C_{yy} w_{2y} = 0$$

(9)

$$w_{2x}^T C_{xx} w_{2x} = w_{2y}^T C_{yy} w_{2y} = 1$$

When the constraint condition formula (9) is satisfied, the correlation coefficient $\rho(w_{2x}, w_{2y})$ between the variables $a_2 = w_{2x}^T x$ and $b_2 = w_{2y}^T y$ reaches a maximum (resolving the constrained optimization problem.

$$\max_{w_{2x},w_{2y}} \rho(w_{2x},w_{2y}) = w_{2x}^T C_{xy} w_{2y} \quad s.t \begin{cases} w_{2x}^T C_{xx} w_{2x} = 1, w_{2y}^T C_{yy} w_{2y} = 1 \\ w_{1x}^T C_{xx} w_{2x} = 0, w_{1y}^T C_{yy} w_{2y} = 0 \end{cases} \tag{10}$$

Among them, $E[a_1 a_2^T] = E[b_1 b_2^T] = 0$ ($w_{1x}^T C_{xx} w_{2x} = w_{1y}^T C_{yy} w_{2y} = 0$) explains that a_1 and a_2, b_1 and b_2 both are linearly independent. Deriving from this, the vector (w_{1x}, w_{1y}) that has been found is used as the first pair of projection vectors, and the combination (ρ_1, a_1, b_1) is obtained. After the $k-1$ sets of projection vectors are determined, we will get a pair of vectors (w_{kx}, w_{ky}) after the k the pair of projection vectors satisfies the constraint formula (11) makes the objective function (1) reach a maximum value.

$$\begin{cases} w_{kx}^T C_{xx} w_{kx} = 1, w_{ky}^T C_{yy} w_{ky} = 1, \\ w_{ix}^T C_{xx} w_{kx} = 0, w_{iy}^T C_{yy} w_{ky} = 0, \\ \quad (i = 1, ..., k-1). \end{cases} \tag{11}$$

For the obtained pair of projection vectors (w_{kx}, w_{ky}), there is always a corresponding combination (ρ_k, a_k, b_k). Where ρ_k is the correlation coefficient between $a_k = w_{kx}^T x_i$ and $b_k = w_{ky}^T y_i$ that has been obtained.

With the above calculations, we can get a series of combinations:

$$(w_{1x}, w_{1y}, \rho_1, a_1, b_1), (w_{2x}, w_{2y}, \rho_2, a_2, b_2), \cdots\cdots, (w_{rx}, w_{ry}, \rho_r, a_r, b_r)(1 \leq r).$$

They have the following properties:

The correlation coefficient gradually decreases. $\rho_1 \geq \rho_2 \geq ... \geq \rho_r$

The projections of x and y on $w_{ix}, w_{iy} (i = 1, 2, ..., r)$ are a_i and b_i. Its variance is

$$D(a_i) = D(b_i) = 1 \ (w_{ix}^T C_{xx} w_{ix} = w_{iy}^T C_{yy} w_{iy} = 1).$$

The projection a_i, a_j of x to w_{ix} and $w_{jx} (i \neq j)$ are linearly independent ($E[a_i a_j^T] = w_{ix}^T C_{xx} w_{jx} = 0$). Similarly, the projection b_i, b_j of y on w_{ix} and $w_{jx} (i \neq j)$ are linearly independent ($E[b_i b_j^T] = w_{iy}^T C_{yy} w_{jy} = 0$).

For the combination $(w_{ix}, w_{iy}, \rho_i, a_i, b_i)$, we call (w_{ix}, w_{iy}) the i th representative projection vector, ρ_i the i th typical correlation coefficient, and (a_i, b_i) the i th representative variable.

3. Simulation Research

In this paper, three kinds of linear frequency modulation signals with different frequency ranges are used. Two or two groups are used as sample signal source for simulation experiments. The specific parameters are as follows:

Signal A: frequency range 50Hz to 450Hz, sampling rate 2000Hz, plus white noise;
Signal B: frequency range 250Hz to 650Hz, sampling rate 2000Hz, plus white noise;
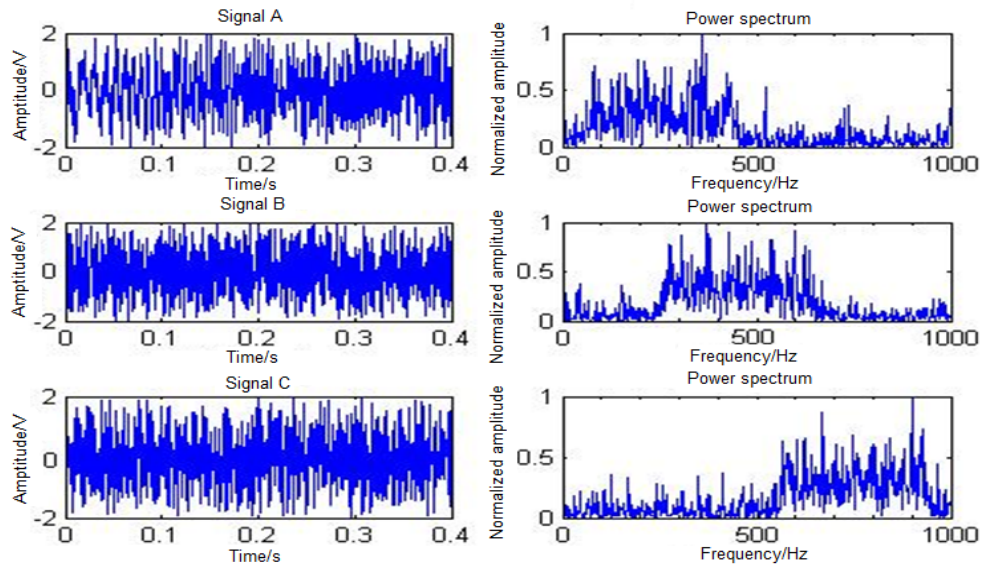Signal C: frequency range 550Hz to 950Hz, sampling rate 2000Hz, plus white noise.

Fig 3 Simulation signal

Figure 3 shows the time domain waveform and spectrogram of the simulated signal.Seen from the figure, the frequency range of signal B overlaps the frequency ranges of signal A and signal C, and signal B and signal A have a large degree of overlap.The next simulation will verify the effectiveness of the algorithm under different frequency overlap conditions.
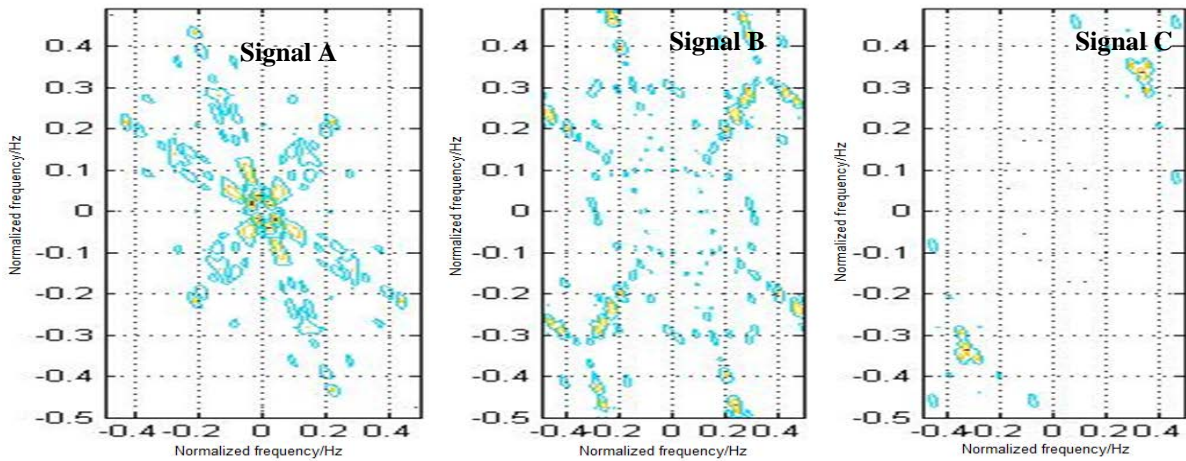

Fig 4 Bispectral features of the simulated signal

Figure 4 shows the bispectrum features of the simulated signal.As can be seen from the figure, there are differences in the bispectrum of the three types of signals.However, this differentiation is not as obvious as the power spectrum characteristics and the Wigner distribution characteristics.This difference is due to the addition of white noise.From Section 2.1 we can see that the bispectrum calculation method is the accumulation of the delay error.In this experiment, the three types of samples selected were 1024 points of data, 32 points per segment, and a total of 32 segments (bispectrum is the average of 32 segments of data).When the data sample is short, the total number of segments decreases (the number of accumulated segments becomes smaller).This also indirectly causes the final accumulation to fail to cancel the white noise in the original signal.For a specific sample in each type of sample, it does not produce a stable bispectrum.This leads to the fact that the resulting bispectrum is not a stable bispectrum, which is detrimental to recognition.

Simulation example: extracting power spectrum features and bispectral distribution-singular value characteristics as two types of sample features.There are three types of situations at this time: a. Signal A and signal C as sample signals; b. Signal A and signal B as sample signals; c. Signal B and signal C as sample signals.

Under the above three types of simulation settings, the canonical correlation analysis method was used for processing and ten experiments were repeated for each case.The recognition rate of each algorithm is shown in the figure below.
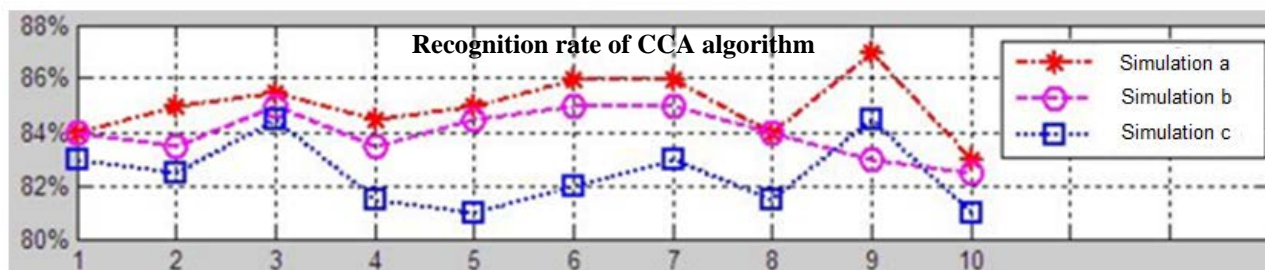
Fig 5 Typical correlation analysis method recognition rate

As can be seen from the figure, the recognition rates obtained in the three cases are all higher than 80%.The recognition rate of the algorithm in class a is always higher than that in the case of class b and c.In category a, signal A and signal C are selected as sample signals, and there is no additive component in the frequency range between the two types of signals.The frequency range of signal B and signals A and C are superimposed (the difference between signal A and signal C is the largest, and these two signals are most easily distinguished and the simulation results are consistent with it).

4. Conclusions

Through the above-mentioned theoretical simulation analysis, the following basic conclusions can be obtained as follows:

(1)The typical correlation analysis method is an effective feature recognition algorithm that can be used to identify the target signal.

(2)For different signals and extracting different signal characteristics, this will have an impact on the performance of the typical correlation analysis method and its variants.We should choose different features according to local conditions to complete the target recognition.

References

[1]Richard J. Vacan's Canonical Correlation Analysis Method, ro(edited). The Past, Present, and Future of Underwater Acoustic Signal Processing. IEEE Signal Processing Magazine, July. 1998 P21.

[2]Bian Yingqi, Zhang Xuegong. Pattern Recognition. Beijing: Tsinghua University Press, 2000.

[3]Cai Yuanlong. Pattern Recognition. Xi'an: Northwestern Telecommunication Engineering Institute Press, 1986.

[4]Bingnan Pei，Zheng Bao，Mengdao Xing.Logarithm bispectrum based approach to radar range profile for automatic target recognition. The Journal of the pattern recognition.2002，35(11)：2643-2651P.

[5]Zhang Yuting, Fang Kaitai. Introduction to Multivariate Statistical Analysis. Beijing, Science Press, 1999.

[6]Xu Xingzhong.Canonical Correlation Coefficients and Canonical Correlation Variables of Multiple Sets of Variables.Science Bulletin, 1996,41(13):1153-1157.

[7]M. Hasan. A new approach for computing canonical correlations and coordinates. International Symposium on Circuits and Systems, 2004, 3: 309 -312.

[8]J. Yang, A. F. Frangi, J. Y. Yang, et al. KPCA plus LDA: a complete kernel Fisher discriminant framework for feature extraction and recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence，2005，27(2):230-44.

[9]Quan-Sen Sun, Sheng-Gen Zeng, Yang Liu, et at. A new method of feature fusion and its application in image recognition. Pattern Recognition, 2005, 38:2437-2448.