2014

ISSN : 0974 - 7435 Volume 10 Issue 22

BioTechnology
An Indian Journal

FULL PAPER

BTAIJ, 10(22), 2014 [13745-13750]

Research on radio signal feature extraction and classification method based on principal component analysis

Liu Yang
Chengdu Radio Monitoring Station of SRMC, Chengdu, 610000, (CHINA)

ABSTRACT

Due to the fast development of technology, especially the rapid changes in radio technology, people's lives are increasingly depending on radio technology. Radio spectrum enriches people's lives, which makes spectrum resources particularly precious. However, at present, there is a lot of abuse of the spectrum in people's lives, which not only leads to a substantial waste of spectrum resources, also sometimes poses a threat to human health. Further, this will cause disruptions to the whole communications industry and thus restrict the rapid socio-economic development. In order to solve this thorny problem, well-known experts and scholars began to study radio signal feature extraction and classification methods. In recent years, many scholars have been trying to settle the issue (namely abnormal signal analysis) in the area of communication by using fuzzy set theory and neural network, and made great achievements. This paper adopts principal component analysis method to extract the features of radio signals, and improves the previous extraction methods by combining entropy method with fuzzy c-means, and applies weighting method to label the importance of each feature in order to solve problem of uneven contribution in classification. It is proved that the improved method is of high efficiency, high speed, capable of determining categories of abnormal signals accurately and efficiently so as to guarantee the integrity and reliability of the information in the communication link, with great practical value.

KEYWORDS

Principal component analysis; Radio signal feature extraction; Radio signal classification method; Fuzzy C-means.

INTRODUCTION

With the rapid development of radio industry, there is an increasing contradiction between supply and demand of radio spectrum resources, and the increasing signal interference has seriously affected the social stability, as well as citizens' lives. Stochastic resonance particularly bi-stable stochastic resonance has unique advantages in amplifying and identifying weak signals[1]. Generally weak signal is interfered by strong noise, in which case, by putting the signal on the stochastic resonance, with the coordination of nonlinear system, the energy transfer happens between signal and noise, and in the process of transfer, there will result resonance output, thus the noise ratio of the system gets improved, and ultimately the goal of weak signal detection is achieved[2]. In the research on stochastic resonance system by Mcnamara B and others, it is found that the change curve of signal-noise ratio varying with the intensity of the noise is the non-monotonic bell-shaped curve, and can be used as an evaluation to measure the system's resonance. After the proposition of this conclusion, signal-noise ratio becomes the most commonly used measure of judgment in the study of resonance vibration of stochastic resonance system. In signal detection, the ratio of output signal-noise ratio to input signal-noise ratio is defined as Gain of Signal to Noise Ratio[3]. In stochastic resonance system, the GSNR can be over 1, and this was proofed by Loerincz K et al. Later people began to be greatly interested in the part of GSNR beyond 1. This study proposes a new classification and analysis method of abnormal signal based on Fuzzy C-means (FCM) and Principal Component Analysis (PCA), and explains its validity by experiences.

Some non-standard irrational use of radio resources affects the normal operations of the radio spectrum resource and accuracy of information transmission. With the development of stochastic resonance system technology, single bi-stable stochastic resonance technology was introduced in the study of weak signal amplifying identification by a lot of experts and scholars, greatly improving the system's GSNR and providing new research approach for detecting weak signal. How to further improve the recognition ability in weak signal detection has now become scholars' direction[4]. Chapeau F et al constructed array stochastic resonance model in 2004. This model is one of saturation threshold, in this model, with the effect of array-noise technology, array GSNR is over 1 compared with single GSNR. Namely after inputting signal into the array stochastic resonance system, array noise can draw out more information and is positive in signal extraction, with a maximum output. Array bi-stable stochastic resonance model was proposed by Duan F et al in 2006[5], and the primary theoretical analysis was processed with its result published in 2008. And the model diagram is shown as Figure 1. This study introduces array bi-stable stochastic resonance model in weak signal detection, combines theory and numerical simulation, and further analyzes GSNR in array bi-stable stochastic resonance system. There is great theoretical significance and practical value in automatic analysis of abnormal signals using new data mining technology with the high processing speed of computer.

DETECTION PRINCIPLES AND ALGORITHM

In detection of C-band signal, substrate noise is the normal signal, and array is with nonlinear characteristic, but this characteristic is inseparable from interaction of array noise and outside noise, so in previous studies, theoretically there is great difficulty to analyze the auto-covariance and cross-covariance function, but the relationship between the two can be analyzed by means of numerical simulation. Preset the signal amplitude, with a value of 0.1, and signal frequency and frequency band width, with the same value of 0.01Hz, sampling rate of 5Hz, and the system's two real parameters set to 1. After the initial value setting, use the model to simulate and get the simulating results, which are change curve of auto-covariance and cross-covariance with array noise intensity, and change curve of auto-covariance and cross-covariance with array external noise. Curve ① shows the variation law of auto-covariance with the change of array noise intensity when the input value of array noise is 0.1[6]; curve ② shows the variation law of cross-covariance with the change of array noise intensity when the input value of external noise is 0.1. Curve ① shows the variation law of auto-covariance with the change of array external noise intensity when the input value of array external noise is 0.1; curve ② shows the variation law of cross-covariance with the change of array external noise intensity when the input value of external noise is 0.1, thus to achieve the objective of dimensionality. Then classify the signal using Fuzzy C-Means clustering analysis method. Fuzzy ISODATA clustering analyzing algorithm was proposed by J.C.Bezdek in 1977. Array bi-stable stochastic resonance model has been mentioned. And the signal is non-coupled parallel. Weak sine signal is denoted as $s(t)$, external noise signal is denoted as $\zeta(t)$, mixed signal is formed by both of them, so is dented as $s(t)+\zeta(t)$, $s(t)$ and $\zeta(t)$ are irrelevant. The input of each array element is mixed signal. In the formula, $s(t)$ can be calculated by $A\cos(2\pi ft)$, among which A represents signal amplitude; the average value of white noise is 0, density is represented by $D\zeta$, namely $\zeta(t)$. Array internal noise, also called array noise, is denoted as $\eta i(t)$, and this noise is irrelevant with mixed signal too. The density of multi array noise is represented by $D\eta$, among which array noise is independent to each other. Each stochastic resonance element in the array has a responding equation. In the above formula, there are two real parameters in the array element stochastic resonance, which is a and b respectively. The output signal of array element is represented by $x(t)$, and the number i element is denoted as $x_i(t)$, the arithmetic mean value of each elements' output is the output of array bi-stable stochastic resonance. The basic idea is to minimize the value of objective function $J(R,V)=\sum_{k=1}^{n}\sum_{i=1}^{c}(r_i k)q\left\|u_k-V_i\right\|^2$ through fuzzy ISODATA clustering analyzing algorithm, thus to realize the optimal dividing. In the formula, $u_k=\{u_{k1},u_{k2},...\}$ represents the number K sample. $(k=1,2,3,...,n)$, $V_i=(v_{i1},v_{i2},v_{i3},...,v_{im})$ $(i=1,2,...,c)$ represents the number I cluster centers. The calculation process is a process of repeatedly modifying the cluster centers and classification matrix. According to the theory and the relationship between auto-covariance and cross-covariance showed in Figure 2, it is described that in the array noise, if the noise is 0, then the auto-covariance and cross-covariance are the same. In array external noise, if the noise is 0, the noise function of each array element is irrelevant. Figure 1 shows the cross-covariance, from which it can be told that with unchanged value of external noise, the cross-covariance reduces while array noise increases, namely away from auto-covariance, which is similar with the inverse relation. Figure 2 tells that with unchanged array noise, the cross-covariance increases while array external noise increases, and gets closer to auto-covariance. Based on the above analysis, the cross-covariance calculation formula can be achieved as shown in Figure 3. Using this formula, in the case of changes of array internal noise and array external noise, the value of cross-covariance can be calculated and its changing regularity can be mapped. Array internal noise is denoted by curve □, from which it can be easily told that this curve is basically identical with the numerical simulated change curve of cross-covariance. Array external noise is also denoted by curve □, similarly, the curve is basically identical with the numerical simulated change curve of cross-covariance. And this verifies the rationality of assumption in cross-covariance formula. By putting the assumed formula into the former formula, it is easy to have an expression for GSNR. Feature extraction is necessary before using fuzzy c-means clustering method for classification and identification of abnormal signal, and the methods to carry out feature extraction directly affect the result of the classification. The GSNR of array element is represented by G1,as shown in Figure 3, array GSNR is not only connected with the intensity of array internal noise, also inseparable with the intensity of external noise, and benefits from the number of array elements and the changes of elements' GSNR. Compared with single GSNR, the fact that value of array GSNR is over 1 is reasonable (Figure 4). This study proposes a method of signal feature extraction based on the fuzzy c-means clustering analysis and principal component analysis to ensure the integrity of the signal feature.

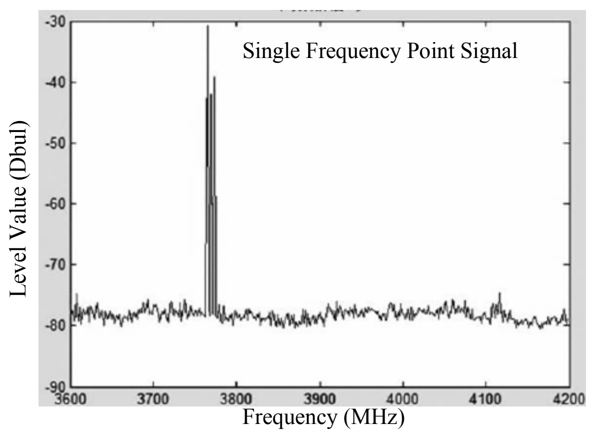

Figure 1 : Single frequency point signal spectrum

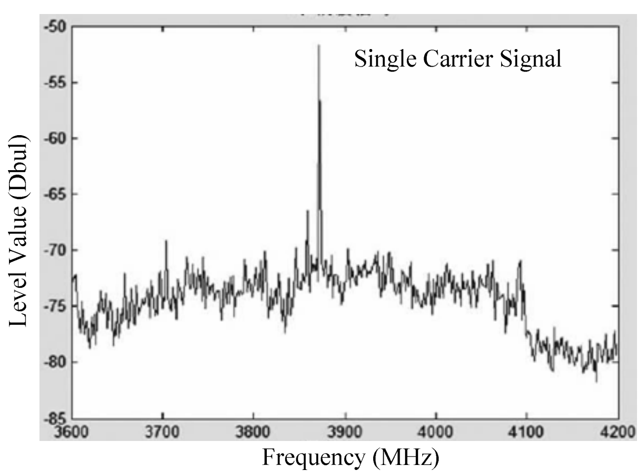

Figure 2 : Single carrier signal spectrum

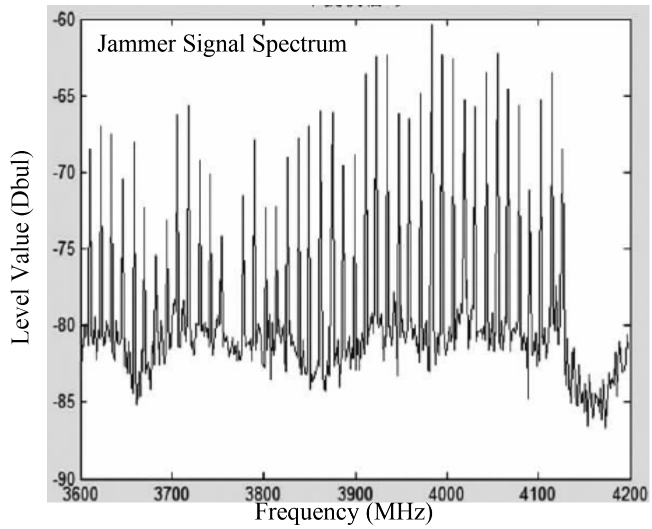

Figure 3 : Jammer signal spectrum

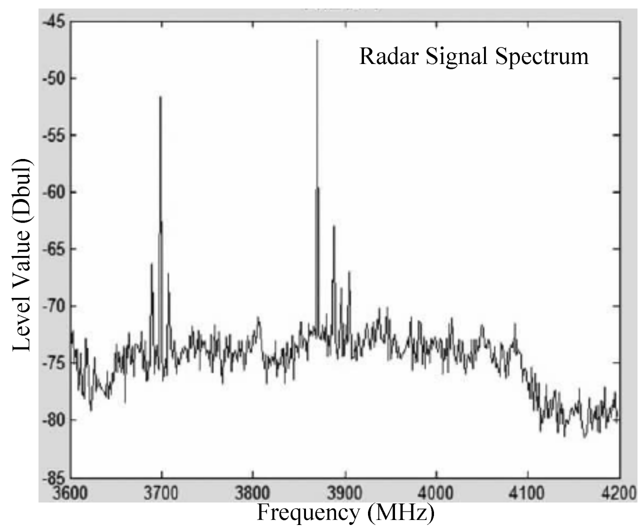

Figure 4 : Radar signal spectrum

Spectrum analyzer is used for collection of four kinds of measured radio abnormal signal (radar, jammer, single carrier and single frequency) for c-band. As for array bi-stable stochastic resonance system, in the detection of weak signal, for full research in its property[7], this study studies the influence of array internal noise, array element number, and array external noise on the array signal-noise ratio respectively[7]. Similarly, preset the signal amplitude as 1, the signal frequency and frequency band width as 0.01Hz, sampling frequency 5Hz, and the two real parameters 1. The change curve of array GSNR along with the intensity of array noise is mapped, as well as the changing trend of array GSNR when the intensity of external noise is changed. There are total five cases of external noise listed. From the map, it can be seen that regardless of the external noise changes, array GSNR can always reach its maximum peak via array noise tuning, and this peak is also the peak of array stochastic resonance. There are 400 sets as test sampling sets. And the algorithm is described as below.

(1) Standardize the sample matrixes:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sqrt{s_j}}, i = 1, 2, 3, \ldots 4m, 4m+1; \; j = 1, 2, \ldots, n$$

among which $\bar{x}_j = \frac{\sum_{i=1}^{4m+1} x_{ij}}{4m+1}$, $s_j = \frac{\sum_{i=1}^{4m+1} (x_{ij} - \bar{x}_j)^2}{4m+1}$, thus there gains a standard matrix $Z = (z_{ij})_{(4m+1)\times n}$. According to the signal type, Matrix Z is divided into $A^{(1)}$, $B^{(1)}$, $C^{(1)}$, $D^{(1)}$, $E^{(1)}$ which represents radar signal, jammer signal, single carrier signal, single frequency point signal, and abnormal signal respectively.

(2) Calculate the correlation matrix of standardization matrix $A^{(1)}$

$$R=\frac{1}{m}A^{(1)}TA^{(1)}$$

(3) Calculate the k-th eigenvalue of the correlation matrix R, then arrange them in descending order. Here the singular value decomposition (SVD) can be used to solve the eigenvalue, eigenvector.

Make $P=\frac{1}{\sqrt{m}}A^{(1)}, R=P^TP$, by using SVD, then there is $P=U\begin{bmatrix} d1 & & & & \\ & d2 & & & \\ & & \dots & & \\ & & & dk & \\ & & & & o \end{bmatrix}V^T$ among which,

$\lambda_1=d_1^2\lambda_2=d_2^2\dots\lambda_k=d_k^2, V$ column vector $(V_1,V_2,\dots V_n)$ is the feature vector of matrix R.

(4) Determine the number of principal components: Determine the value of R by $\frac{\sum_{i=1}^{r}\lambda_i}{\sum_{i=1}^{k}\lambda_i}\geq 0.90$ can makes the information utilization ratio up to above 90%, using $Rb=\lambda_j b$ can calculate the unit eigenvector, $b_j\,(j=1,2,3,\dots r)$ and different feature vectors are mutually orthogonal.

(5) Determine the principal components: Feature vector $b_j\,(j=1,2,3,\dots r)$ makes up matrix of $B_{n\times r}$, $C=(C_1,C_2,C_3,\dots C_r)=A_{m\times n}^{(1)}B_{n\times r}$, $C_1,C_2,C_3,\dots C_r$ are the extracted principal components. 3 Repeat the step 2 to different type signals of $B^{(1)}, C^{(1)}, D^{(1)}$, then there gets the matrixes composed of each signal principal components $C_{m\times r}^{(i)}(i=1,2,3,4)$.

(6) Clustering. Using FCM to cluster each of the above processed matrixes respectively, $C_{m\times r}^{(i)}(i=1,2,3,4)$ and there will be $k(i)$ clustering centers.

(7) Process new signals.

Multiply new signal $E(1)$ with matrix, extract the principal components $c_0^{(i)}=(c_{o1},c_{o2},c_{o3},\dots c_{or})$

Compare with the clustering centers in step 4, and calculate the distance between these two. Clarify the new signal to signal type with minimal distance.

(8) The experiment is repeated for all kinds of signals, calculating the recognition rate.

ANALYSES OF EXPERIMENTAL RESULTS

Set the signal amplitude as 1, the signal frequency and frequency band width as 0.01Hz, sampling frequency 5Hz, and the two real parameters 1. When the number of array element is changed, the change curve of array GSNR varying with the intensity of array noise is shown in Figure 4. From Figure 4, it can be seen that when the number of array element increases, the peak of ratio between the array GSNR and the array elements' GSNR increases too, which means compared with single stochastic resonance, array stochastic resonance's performance for weak signal detection is significantly improved. Notably, when the number of array element is less than 100, its performance is significantly improved, but when the number of array element is large enough, its performance improvement is not satisfactory. This study uses principal component analysis method and fuzzy c-means for experimental operations on training samples and test samples to achieve automatic recognition result. Analysis of array bi-stable stochastic resonance model and GSNR has been made in the above part of this paper, in order to proof the detection performance of this model in weak signal, weak periodic signal is chosen for simulation test. And the test result is shown in TABLE 1.

TABLE 1 : Recognition effects based on principal component analysis method and fuzzy C-means clustering signal classification

	Radar	Jammer	Single Carrier	Single Frequency Point
Number of Principal Component	5	4	4	4
Accuracy Rating	1	0.925	0.8	0.825

TABLE 2 is mean values based on experts' experience. In the experiment, when it comes to the set of parameters of the model, the signal amplitude is set to 1, the signal frequency and the bandwidth value is 0.01Hz, the sampling frequency is set to 5Hz, two real parameters of the system are set to 1, the number of array elements is set to 100, and intensity of array noise is set to 0.2. The third peak is the experimental accuracy of the input index.

TABLE 2 : Experimental feature extraction result based on the experiences of experts

	Radar	Jammer	Single Carrier	Single Frequency Point
Accuracy Rating	0.91	0.873	0.76	0.725

The test results show that the ratio of array GSNR is over 1, which means the array GSNR is obviously large. Compared with the single stochastic resonance detection, array bi-stable stochastic resonance performs better in detection of weak signal. But if the external noise increase, the detection performance of bi-stable stochastic resonance will be reduced, and the bigger the external noise is, the performance of bi-stable stochastic resonance is more similar with the single one. And this result is completely consistent with the theoretical analysis.

CONCLUSION

This study adopts principal component analysis method to extract the features of radio signals, and improves the previous extraction methods by combining entropy method with fuzzy c-means, and applies weighting method to label the importance of each feature in order to solve problem of uneven contribution in classification. In recent years, many scholars have been trying to settle the issue of abnormal signal analysis in the area of communication by using fuzzy set theory and neural network, and made great achievements.

REFERENCE

[1] Kang Yuan, Li Jinghua; Principal component feature extraction technology based on time-frequency analysis for battlefield acoustics signal [J], Electronic Measurement Technology, **25(9)**, 37-41 **(2013)**.

[2] Ren Jing-Bo, Sun Genzheng, Chen Bing, Luo Ming; Milling chatter identification method based onwavelet packet transform and kernel principal component analysis[J], Noise and Vibration Control, **20(2)**, 67-72 **(2014)**.

[3] Zhang Jianhua, Qi Lijun, Ji Ronghua, Yuan Xue, Li Hui; Classification of cotton blind stinkbug based on gabor wavelet and color moments[J], Transactions of the Chinese Society of Agricultural Engineering, **30(6)**, 67-74 **(2014)**.

[4] Wang Qianqian, Qin Keyun; Classification MTHOD of C band radio signal based on PCA and FCM[J], Journal of Minnan Normal University (Natural Science), **30(6)**, 67-74 **(2014)**.

[5] Ye Yongwei, Liu Zhihao, Huang Liqun; Signal feature extraction for automobile coating equipment based on PCA[J], Chinese Journal of Scientific Instrument, **15(6)**, 32-36**(2013)**.

[6] Xiao Zhi, Li Wenjuan; Personal credit scoring based on PCA and SVM [J], Technology Economics, **30(5)**, 131-136 **(2012)**.

[7] Liu Ping; Research on face recognition methods based on entropy-based local preserving and statistical feature extraction[J], Youth Literator, **20(12)**, 131-134 **(2013)**.