Aeronautical $\gamma$ Spectrum Noise Reduction Method Based on LS-SVM Segmentation Regression

Ying Ci$^1$, Jun Liu$^2$, Lei Chen$^2$, Guang Li$^2$, Feng Yang$^2$

$^1$School of Information Science and Engineering, Northeastern University, Shenyang, Liaoning 110819, China
$^2$Engineering Research Center of Nuclear Technology Application (East China University of Technology), Ministry of Education, Nanchang, Jiangxi 330013, China

*Corresponding author’s e-mail: jun_liu_2003@163.com

Abstract: In order to reduce noise of the aeronautical $\gamma$ spectrum caused by factors such as less counting, short measuring time and poor measuring environment, this article adopts a segmented noise reduction method combined with the machine learning method, LS-SVM (Least Squares Support Vector Machines) and weighted stacking method according to the energy window distribution of the spectrum. The experimental results show that the segmented noise reduction method based on LS-SVM can significantly reduce the statistical fluctuation. In addition, the method presents good adaptability and generalization ability.

1. Introduction

Aeronautical $\gamma$ detection is a fast, economical and efficient method for regional nuclear geophysical exploration\cite{1}. Aeronautical $\gamma$ spectrum generally uses Three-window method proposed by International Atomic Energy Agency (IAEA) to analyze the content of Th, U and K in the stratum. The Three-window method relies on counting in the energy window, and the result is affected by noise in the energy window\cite{2}. Therefore, the key step in analyzing the aeronautical $\gamma$ spectrum is noise reduction to the spectrum.

Guo et al. proposed least squares moving smoothing method, which is applied to the aeronautical $\gamma$ energy spectrum denoising. The selection of the fitting points depends manually, but the adaptability is poor\cite{3}. Zhang’s team proposed a wavelet denoising method that can reduce high-frequency noise with a very small change to the spectral shape after smoothing. However, it is necessary to select different decomposition reconstruction scales manually for processing different energy spectrum. Thus, this method is not suitable for processing aeronautical $\gamma$ analysis with large amount of data\cite{4}.

The noise adjusted singular value decomposition method (NASVD) and the maximum noise fraction method (MNF) are the most widely used methods for aeronautical $\gamma$ spectrum denoising, both of which extract mutually orthogonal principal components of the spectrum from the same series of the measured spectrum, and arrange them in descending order according to the contribution degree to the spectral shape. The low-order part is the main composition of the spectral shape, and the higher one is noise. The low-order part is reconstructed to achieve noise reduction. But these two methods require a large number of spectra of the same series, while they are greatly affected by the actual working conditions.
2. LS - SVM Noise Reduction Method

2.1. Principle

Support Vector Machine (SVM) is a machine learning method proposed by Corinna Cortes and Vapnik in 1995[7]. In small sample, nonlinear and high dimensional pattern recognition, SVM embodies the advantages of fast learning and avoiding local extreme points. It also has the advantages of strong generalization ability and small structural risk. For the denoising processing of aeronautical γ spectrum, which short measurement time, less channel and counting leads to less training sample sets and large feature differences. Due to the large empirical error of the training samples, excessive reduction of training errors leads to over-fitting and reduces generalization[8]. In the denoising process, it is necessary to retain the actual information of the spectrum as much as possible. Hence the noise reduction algorithm not only needs to reduce the training error, but also needs a good generalization.

Based on LS-SVM, the energy spectrum regression fitting model is established, and the measured γ energy spectrum data \((x_i, y_i)\) is given as a training set. where \(i = 1, 2, \cdots, n\), \(x_i\) is the channel address, \(y_i\) is the counting of channel \(i\). The energy spectrum data is mapped to a high-dimensional feature space, and transform the nonlinear data to a linear form in the high-dimensional feature space. Then the energy spectrum regression function is constructed in the high-dimensional feature space. In this article, the radial basis kernel function (RBF) is chosen to simplify the mapping process of energy spectrum data. The mapping process is shown in equation (1).

\[
\omega^T \phi (x_i) + b
\]

LS-SVM transforms the inequality constraints in the traditional SVM optimization problem into equality constraints, as shown in equation (2).

\[
\begin{align*}
\min_{\omega, \xi} J(\omega, \xi) &= \frac{1}{2} \|\omega\|^2 + \frac{C}{2} \sum_{i=1}^{n} \xi_i^2 \\
\text{s.t.} \quad y_i = \omega^T \phi (x_i) + b + \xi_i, i = 1, \cdots, n 
\end{align*}
\]

Where, \(C\) and \(\xi\) are determined by cross-validation, and the Lagrange function is introduced to convert equation (2) into a dual problem, that is:

\[
\begin{align*}
L(\alpha, \xi, \omega) &= J(\alpha, \xi) - \sum_{i=1}^{n} \alpha_i \left( \omega^T \phi (x_i) + b + \xi_i - y_i \right)
\end{align*}
\]

Solving equation (3), the energy spectrum regression curve is given in equation (4).

\[
f(x_i) = \sum_{i=1}^{n} \alpha_i K(x, x_i) + b
\]

And the RBF kernel function is shown in equation (5).

\[
K(x, x^T) = \exp \left[ -\frac{|x - x_i|^2}{\sigma^2} \right]
\]

2.2. LS-SVM aeronautical γ spectrum full-spectrum regression fitting and analysis

The LS-SVM regression hyperplane is linear in the high-dimensional space, and full-spectrum distribution characteristics is shown in the hyperplane[9]. Figure 1 shows the LS-SVM fitting denoising results. As shown in figure 1, the full-spectrum regression fit curve of LS-SVM fitting method can not completely fit the peak shape in channel addresses 110-130. There is still so much noise in the curves of the K and Th energy windows.
The factors of the failure to a good denoising effect by full-spectrum regression fitting algorithm are analyzed as follows:

1. The difference of the channel address and counting between the low-energy segment and the high-energy segment of the energy spectrum are large.
2. The channel address of the aeronautical spectrum is less.
3. In the high-energy segment, there is low counting with large relatively noise.

These factors lead to too many support vectors for LS-SVM full-spectrum fitting process which result in over-fitting phenomenon, so it does not meet the purpose of reducing noise very well. Therefore, in this article, we propose LS-SVM segmentation regression fitting method to train data according to energy window based on the energy distribution characteristics of the aeronautical γ spectrum.

3. LS-SVM segmentation noise reduction method

3.1. Principle of Three-window method

Three-window method is commonly used to aeronautical γ spectrum analysis. The energy window distribution represents the range of each segment of the γ energy spectrum. Segment TC starts with channel address 34, and the channel address 114 is the beginning of the K energy window. Then the following are the three energy window segments. The variation of the counting of each different energy windows is quite different, so we set the segment A (channel addresses 1-34), segment B (channel addresses 35-114), segment C (channel addresses 115-130), segment D (channel addresses 131-154) and segment E (channel addresses 155-256) to perform segmental observations on 20s sets of the measured aeronautical γ spectrum. The standard deviation of the energy window counting (the degree of the change of the counting of each energy window) is taken as the characteristic observation parameter of the energy spectrum.

The difference between the standard deviation of the energy windows among the 20s energy spectrum is small, so it can be considered that the counting of the aeronautical γ energy spectrum distributes regularly with the variation of the channel address. According to statistics, the mathematical expectation of the counting standard deviation of energy segment A is 196.43, and which of segment B is 29.36. Obviously, the standard deviation difference between the two segments (A and B) is large, thus the two energy windows can be divided by the standard deviation. While, the mathematical expectation of the counting standard deviation of segment C is 9.76, which of segment D and E are 3.35 and 4.77. Because the standard deviation of the counting among the last three energy windows (C, D and E) is small, We try to set the channel addresses 115-256 as a whole energy segment.

Moreover, in order to avoid excessive change of the first derivative at the segmentation point, which makes the segmentation learning processing result not smooth and makes the data gap to appear at the segmentation point. So this article propose a weighted superposition smoothing processing at the segmentation point based on LS-SVM segmentation learning noise reduction processing, as shown in equation (6).
\[
Y(i) = \begin{cases} 
    y_1(i); & 1 \leq i < 34 - t \\
    y_2(i); & 34 + t < i < 114 - t \\
    y_3(i); & 114 + t < i < 256 \\
    k_j \cdot y_1(i) + \overline{k}_j \cdot y_2(i); & 34 - t \leq i \leq 34 + t \\
    k_j \cdot y_2(i) + \overline{k}_j \cdot y_3(i); & 114 - t \leq i \leq 114 + t 
\end{cases}
\]

Where, \(i\) is the channel address, \(y_1, y_2, y_3\) is fitting regression results of the three spectrum segments. The channel address distributions of the three segments are \((1 \sim 34+t), (34-t \sim 114+t)\) and \((114-t \sim 256)\). Obviously, the channel addresses overlap between each segments is \(2t+1\). \(k_j\) and \(\overline{k}_j\) are smoothing weighting factors, which are shown in equation (7) and (8)**.[10]**

\[
k_j = \frac{1}{2} \frac{1}{2} \sin\left(\frac{j-1}{2t} \pi + \frac{\pi}{2}\right) \quad j = 1, 2, \ldots, 2t+1
\]

(7)

\[
\overline{k}_j = \frac{1}{2} \frac{1}{2} \sin\left(\frac{j-1}{2t} \pi + \frac{\pi}{2}\right) \quad j = 1, 2, \ldots, 2t+1
\]

(8)

By experiments, the count value of the overlapping channel addresses with the best weighted superposition smoothing effect is 7.

4. Experimental results and discussion

The LS-SVM segmentation regression method was used to reduce the noise of 100 aeronautical \(\gamma\) spectrum measured at a fixed point. A comparison between one of the spectrum and the average of the entire spectrum is shown in Figure 3. The average spectrum in figure 2 was considered as the true value of the spectrum.

As shown in Figure 2, it is clear that after the LS-SVM segmentation regression, the spectral line noise was significantly reduced and the peak shape remained good. The counting of the K, U, and Th three-energy windows were gentle, unaffected by noise, and no data faults appeared at the energy spectrum curve segmentation point, which avoided the shortcomings of LS-SVM full-spectrum smooth noise reduction. The measured aeronautical energy spectrum data were compared with the results obtained by the method of 15-points least squares moving method. The data comparison results are shown in Figure 3.

![Figure 2. The result of LS-SVM denoising and average spectrum.](image-url)
Figure 3. Comparison between LS-SVM and smoothing denoising. The smoothing effect is characterized by the parameter $V$ and $D(V)$, as shown in equation (9)\cite{10}.

$$V = \frac{s_i - m_i}{\sqrt{m_i}}, i = 1, \ldots, 256$$

(9)

Where $s_i$, $m_i$ are counting of the smoothed spectrum and the average spectrum at channel address $i$. $V$ is the deviation between the smoothed spectrum and the average spectrum, and it approximates the normal distribution. $D(V)$ is the variance of $V$. When $D(V)$ is smaller the noise reduction effect is better. The noise reduction effects of LS-SVM and least squares moving smoothing method are shown in table 1.

| Noise reduction method | D(V)   |
|------------------------|--------|
| LS-SVM                 | 0.333  |
| 11point                | 54.477 |
| 13point                | 66.877 |
| 15point                | 36.855 |
| 17point                | 40.842 |
| 19point                | 116.489|

Based on results from Figure 3, the 15-point least squares moving smoothing method can achieve a favourable noise reduction effect in the three-energy window segment, but it causes the spectral shape declining and broadening in channel addresses (1-114). While, LS-SVM segmentation regression fitting method in this article approaches a favorable result. And from table 1, it can be concluded that the method in this article reduces the noise better than the least squares moving smoothing method. So, the LS-SVM segmentation regression method meets the requirements better.

5. Conclusion

In this article, the energy spectrum denoising method is studied based on the characteristics of aerial $\gamma$ spectrum. Conclusions can be drawn by analysis of theoretical derivation and experiment as: (1) According to energy window feature, we proposed an energy spectrum denoising method based on LS-SVM segmentation regression fitting. The method used self-adaptive learning to reduce noise and maintain the peak shape. (2) The LS-SVM learning noise reduction method in this article has both the local and global unique optimal solution. And the cross-validation method is used to determine the kernel function parameters of each energy window, which shows strong adaptability.

Experiments showed that the smoothing effect in this article on the aeronautical $\gamma$ spectrum achieves 0.333, which is better than the moving smoothing.
Acknowledgment
This article is supported by Natural Science Foundation of Jiangxi Province (No. 20171BAB202028). We thank Jiaxuan Chen, the editor, and the anonymous reviewers for their work improving this manuscript for publication.

References
[1] Wan, J.H., Xiong, S.Q., Fan, Z.G. (2012) Current Situation and Prospect of Aviation Gamma Spectrum Measurement Technology. Geophysical and Geochemical Exploration, 36(3): 386-391.
[2] International Atomic Energy Agency. (2003) Guidelines for radioelement mapping using gamma ray spectrometry data. 52-71.
[3] Guo, Y.F. (2003) Smoothing Filtering Processing of Natural Gamma Spectrum. Journal of Northeast Petroleum University, 27(3): 113-117.
[4] Zhang, J.M., Shi, Q.L., Bai, T. (2005) Effect of Wavelet Analysis on Reducing Statistical Fluctuation of Gamma Spectrum. Atomic Energy Science and Technology, 39(4): 349-353.
[5] B Minty. (2003) Accurate noise reduction for airborne gamma-ray spectrometry. Exploration Geophysics, 34(3): 207-215.
[6] Yang, J., Ge, L.Q., Zhang, Q.X. (2010) Application of NASVD Method in Noise Reduction of Aerial Gamma Spectrum Data. Uranium Geology, 26(2): 108-113.
[7] Li, G.Z., Wang, M., Zeng H.J. (2004) Introduction to Support Vector Machine. Beijing: Publishing House of Electronics Industry.
[8] Suykens, J A K. (2014) A robust least squares support vector machine for regression and classification with noise. Neurocomputing, 140: 41-52.
[9] LIU, J., GUAN, X., WU, H.X., et al. (2016) A Smooth method of γ Spectrum Based on LSSVM. Nuclear Electronics&Detection Technology, 36(2): 197-204.
[10] Lv, Y., Liu, J.Z., Zhao, W.J. (2015) Steady state detection method based on piecewise curve fitting. Chinese Journal of Scientific Instrument, 33(1): 194-200.