Combining 3D Robust Smoothing Filtering and Adaptive Clustering Analysis for Geophysical Signal de-nosing and Image Segmentation

Liu Weiqiang[1, 2, 3]

[1]College of Geophysics, China University of Petroleum-Beijing (CUP), Beijing, 102249, China

[2]Key Laboratory of Geophysical Electromagnetic Probing Technologies ofMinistry of Natural Resources, Institute of Geophysical and Geochemical Exploration, Chinese Academy of Geological Science, Langfang 065000, China

[3]Chinese Academy of Geological Sciences, Beijing, 100037, China

Abstract: ***Image segmentation is widely used to enhance images in many fields. However, the existence of noise interferences will distort the segmentation results. In this paper, we developed ahigh-dimensional robust smoothing filter algorithm based on M-estimation. The algorithm can remove both Gaussian random noise and peak pulse interference inone-dimensional (1D), two-dimensional (2D) and three-dimensional (3D) images. We combinedthe robust smooth filtering algorithm withan adaptive clustering analysis algorithm to segment geophysical image and extract the large-scale boundaries of anomalous bodies. By removing noise interferences and marking outliers of the 3D images in advance, the adaptive clustering results are more stable and the resolution is improved. The results of geophysical image segmentation under different noise interference demonstrate the effectiveness of the combined algorithm.***

Keywords: Geophysical signal processing, signal de-noising, robust smoothing, clustering analysis

1. Introduction

Geophysics is concerned with the application of physical principlesto the study of the Earth. Geophysical results are usually expressed by three-dimensional images, such as seismic gray-scale image, resistivity model image, electric logging image, gravity and magnetic anomaly image, etc. These images reveal how the physical properties of the earth's interior vary vertically and laterally, which is more intuitive and vivid than conventional data [1-7].

In order to identify and analyze the abnormal bodies in the image, relevant regions need to be separated and extracted. On this basis, it is possible to make further use of the target, such as identifying the boundary and central position of the abnormal bodies. Image segmentation is dividing an image into varioussub-areasand extracting the interesting objects. Adaptive clustering analysis algorithm has more advantages than traditional methods in image segmentation. It is an unsupervised machine learning algorithm that automatically clusters images based on statistical characteristics of pixel distribution to trace out image boundaries[8-12].

Geophysical data are usually observed in the field and the resulting images are often contaminated by noise interferences from various artificial and natural sources.The Gaussian-like random background noise will blur the image, and make the boundaries obtained by image segmentation become very tortuous and complicated. Outliers caused by spike interference also have a significant impact on adaptive clustering analysis. A small amount of outliers will distort the clustering results. In geophysical image processing, we mainly aim to obtain large-scale anomalous body boundaries. Therefore, it is necessary to reduce the noise interferences in advance to obtain a stable image segmentation result.The filtering method based on robust statistics can suppress the Gaussian noise and the peak pulse noise simultaneously. Compared with mean filtering and median filtering, it can automatically recognizeoutliers and reduce the weight of them to improve image quality [13-18].

In this paper, a 3D robust smooth filtering algorithm is implemented, which can reduce the noises of 1D, 2D and 3D signals and identify global outliers. The algorithm combined with adaptive clustering algorithm can be used for image segmentation and demarcating the boundaries of geophysical anomalous bodies.The codes and simulated data are available by contacting the corresponding author.

2. Three-Dimensional Robust Smooth Filtering Algorithm

A moving average filter smooths data by replacing each data point with the average of the neighboring data points defined within the span window. For data sequence in a smooth window $[S_1, S_2, S_3, \ldots\ldots, S_N]$, the robust mean is calculated by solve the following equation according to robust M-estimationcriterion [19]:

$$\sum_{i=1}^{N} \psi\left(\frac{S_i - \theta}{\sigma}\right) = 0, \quad i = 1 \sim N, \tag{1}$$

where, θ is the robust mean value to be estimated, σ is a scale parameter representing the residual distribution between observed and true values, ψ is an influence function to weight all data points, which has various forms. The "Hampel" function is adopted in this paper. The solution of the equation (1) is calculated by iterative modified residuals algorithm. Firstly, the median value and median absolute deviation of the originalobservation sequence are taken as the initial estimates:

$$\begin{cases} \theta_0 = \text{median}\{S_i\} \\ \sigma_0 = \text{median}\{|S_i - \theta_0|\} \quad i = 1: \; N \end{cases}, \tag{2}$$

Then, the optimal solution is calculated by the following iterative formula:

$$\theta_{k+1} = \theta_k + \frac{\frac{1}{N}\sum_{i=1}^{N}\psi(\frac{S_i - \theta_k}{\sigma_0})\sigma_0}{\frac{1}{N}\sum_{i=1}^{N}\psi'(\frac{S_i - \theta_k}{\sigma_0})}, \tag{3}$$

where, $\psi'(:)$ is the derivative of the influence function $\psi(:)$. When the $\frac{1}{N}\sum_{i=1}^{N}\psi'(\frac{S_i - \theta_k}{\sigma_0})$ is equal to 0, the denominator should be replaced by a constant C, $\frac{1}{2} < C < 1$. When the relative difference between two iterative solutions is less than 1%, the iteration stops.

For a $N \times M \times L$ 3D image, the pixel value can be expressed as follows:

$$S(i, j, k), \; i = 1 \sim N, \; j = 1 \sim M, \; k = 1 \sim L, \tag{4}$$

we firstly set the smoothing filter window $[x, \; y, \; z]$, andthen calculated the M-estimation value $\bar{S}(i, j, k)$ of the data points in the 3D neighborhood of the original data $S(i, j, k)$, namely,

$$\bar{S}(i, j, k) = \text{M\_estimate}\left\{S(i - \frac{x}{2} : i + \frac{x}{2}, \; j - \frac{y}{2} : j + \frac{y}{2}, \; k - \frac{z}{2} : k + \frac{z}{2})\right\} \tag{5}$$

Since the smoothing calculation of each filter windowis independent, we adopted parallel algorithm to calculate multiple filtering windows simultaneously.Finally, the smooth image is obtained by replacing all the pixels $S(i, j, k)$ in the original image with robust estimated values $\bar{S}(i, j, k)$. The separated noise interferences are also obtained by calculating the difference between the original image and the smoothed image.

$$Noise(i, j, k) = S(i, j, k) - \bar{S}(i, j, k), \quad i = 1 \sim N, \; j = 1 \sim M, \; k = 1 \sim L \tag{6}$$

The separated noise interference undergoes further statistics and analysis. Data thatare larger than the sum of noise mean value and three standard deviations is considered as outliers, and the locations of themare marked.

The above algorithm can process images in 3D dimensions. The 1D and 2D image smoothing can be considered as special cases of 3D image. We demonstrated the effect of smooth filtering and outlier recognition by a simulated 1D signal (Fig. 1). Fig.1a shows a noisy sinusoidal signal interfered by random noises and spike pulse outliers.Fig. 1b shows the real signal and the de-noising signal after robust filtering, the root mean square error of the two signals is 0.034. Fig.1c shows the separated noises by calculating the difference between the noisy signal and the smoothed signal. Data points whose difference from the mean value of noise are greater than three standard deviations and 1.5 standard deviations are identified and marked. Both global and local outliers are identified because the separated noise does not include large scale trend drift.

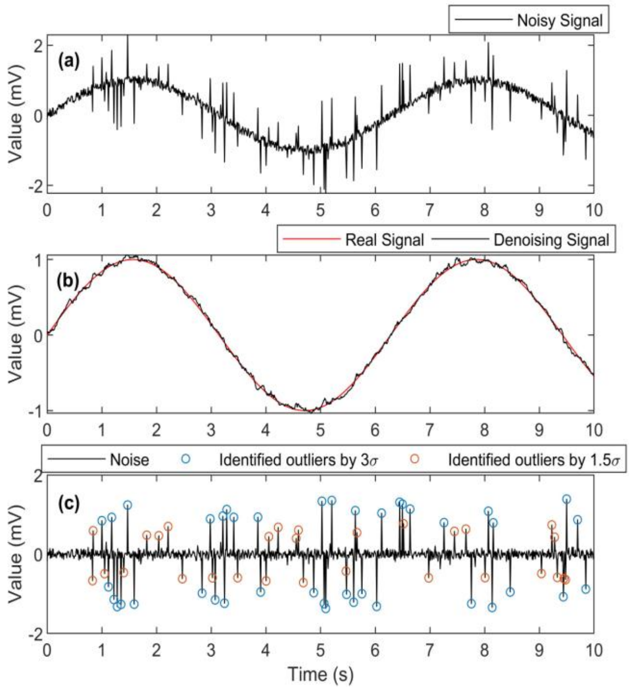

Figure 1: Schematic diagram of robust smooth filtering for de-noising and identifying outliers

3. Robust Filtering and Adaptive Clustering

Adaptive clustering analysis is an unsupervised machine learning algorithm that can automatically cluster images according to the distribution of pixel values, and each cluster has similar properties. Image segmentation and boundary recognition can be achieved by further gradient detection of clustering results. However, the existence of noise interferences will distort the clustering results. Robust smoothing filtering can be used to suppress noise to obtain stable clustering results. The processing is as follows: First, we use the robust filter to smooth all data points in the image, and identify outliers whose amplitude is greater than the sum of mean and three standard deviations, and mark them as a separate category. Then the adaptive clustering algorithm [12]is used to cluster the smoothed data.The algorithm takes the geometric average of the distance between all data points as the threshold, and gets the optimal clustering through multiple iterations. Finally, the data points in the same cluster are assigned the same value and the large scale boundariesare identified by gradient detection of clustering results.

Here, we first illustrate the influence of outlier interference on adaptive clustering by taking scatter data as an example (Fig.2).Fig. 2a shows a set of scatter data which were divided into three clusters. Fig. 2b shows the adaptive clustering results of data points with 2% outliers. The existence of outliers reduces the resolution of the adaptive clustering algorithm, and the original three kinds of detection points are classified into one class. Fig. 2c shows the outliers identified by robust smooth filtering algorithm. Data greater than three standard deviations and less than negative three standard deviations are grouped into two categories. The remaining data is then used for further clustering, Fig. 2d shows theadaptive clustering results after marking outliers. The resolution of the clustering is improved.

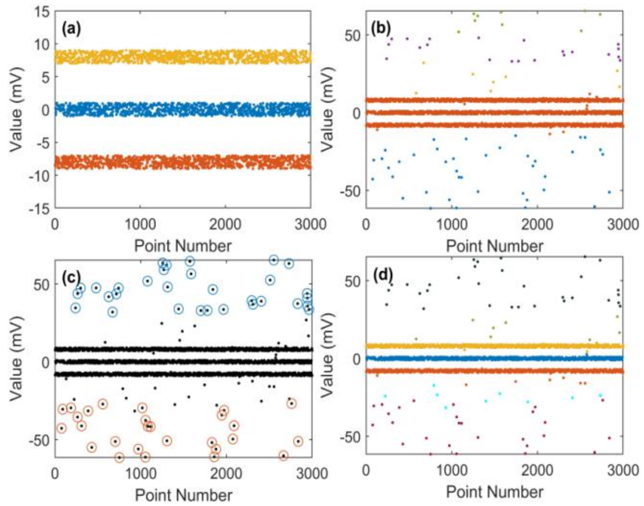

Figure 2: The influence of outliers on adaptive clustering algorithm: (a) adaptive clustering results of data points without outliers, (b) adaptive clustering results of data points with 2% outliers, (c) outliers identified by robust smooth filtering algorithm, (d) adaptive clustering results after marking outliers.

4. Geophysical Imagine De-Noising and Boundary Identification

In order to verify the effect of robust smooth filtering and adaptive clustering analysis, a 3D geophysical resistivity image withvarious noises was used for testing and analysis. The results show that the algorithm can avoid the distortion of clustering result caused by noise interference and improve the recognition accuracy of large-scale boundary.Fig. 3a shows a 3D geophysical image contaminated by 1%scattered outliers caused by impulse noise. Fig. 3b shows the adaptive clustering results of the image. A small number of outliers and a large number of normal points are grouped separately. The large scale boundary of the anomalous body cannot be distinguished by gradient detection. Fig. 3c shows the 3D image after robust smooth filtering. Fig. 3d shows the clustering results. Large scale boundaries identified by gradient detection of clustering results are also marked on the above two graphs. Robust filtering suppressed outlier interference and the resolution of adaptive clustering was improved.

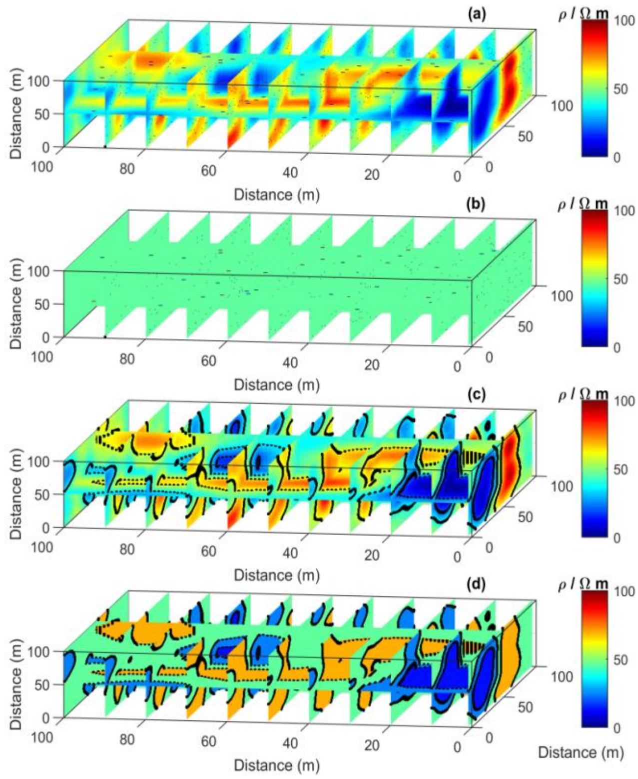

Figure 3: Clustering and boundary recognition of a 3D geophysical image contaminated by 1% scattered outliers: (a) the contaminated 3D image, (d) adaptive clustering and boundary recognition results, (c) the 3D image after robust filtering, (d) clustering and boundary recognition results.

Fig. 4a shows the 3D geophysical image contaminated by Gaussian noise. Gaussian random noise makes the original image very fuzzy. The large scale abnormal bodies are difficult to recognize. Fig. 4b shows the adaptive clustering and boundary recognition results. Although the resolution of the clustering results does not decrease, the clustering results become very scattered, and the corresponding boundaries are also very convoluted and difficult to distinguish. Fig. 4c shows the 3D image after robust smooth filtering. Noise interference is removed and large-scale contours are displayed.Fig. 4d shows the clustering results and boundaries identified by gradient detection.

Fig. 5a shows the 3D geophysical image contaminated by both Gaussian noise and 1% outlier interferences. Fig. 5b shows the adaptive clustering results. Gaussian noise reduces the quality of the original image and outlier interferences distort the clustering results. The boundaries of large-scale anomalous bodies are still difficult to identify. Fig. 5c shows the 3D image after robust smooth filtering. Fig. 5d shows the clustering results and identified boundaries. Both Gaussian noise and outlier interference are effectively suppressed. The results of adaptive clustering and image segmentation are stable and clear.

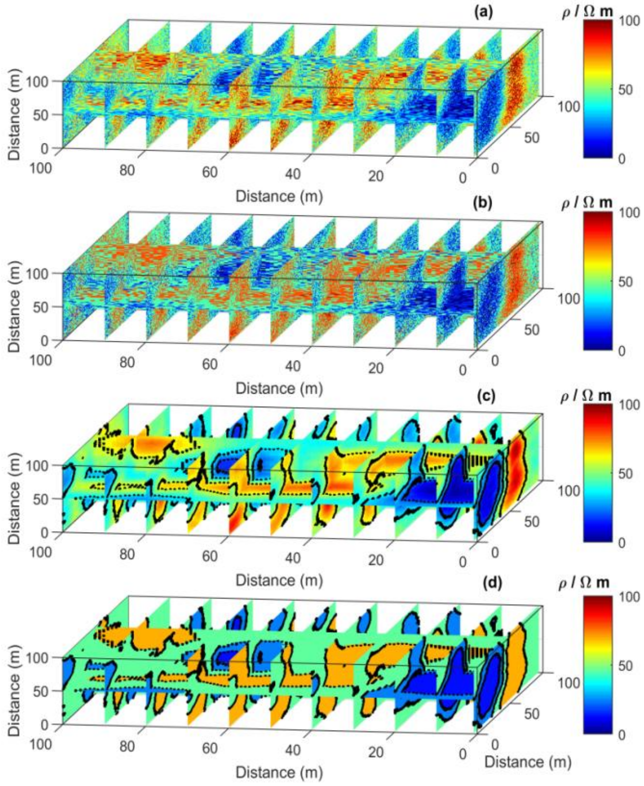

Figure 4: Clustering and boundary recognition of a 3D geophysical image contaminated by Gaussian random noise: (a) the contaminated 3D image, (d) adaptive clustering and boundary recognition results, (c) the 3D image after robust filtering, (d) clustering and boundary recognition results.

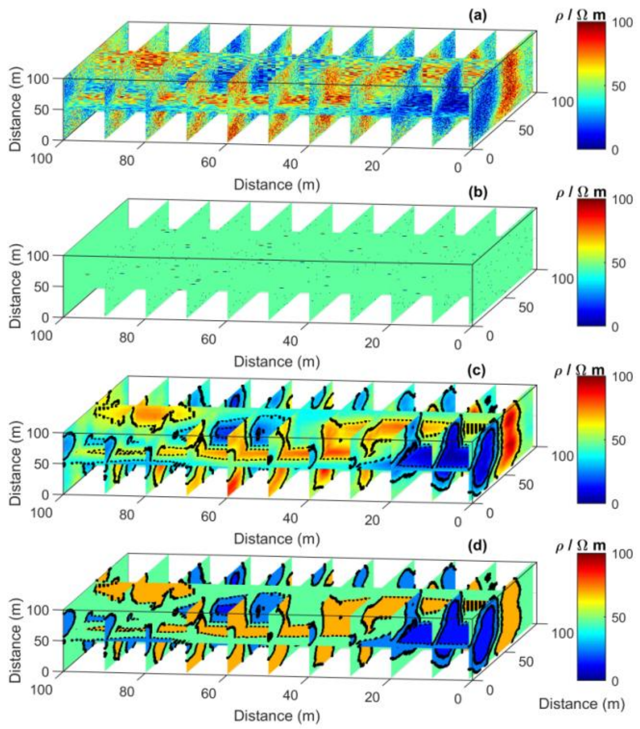

Figure 5: Clustering and boundary recognition of a 3D geophysical image contaminated by bothGaussian random noise and 1% scattered outliers: (a) the contaminated 3D image, (d) adaptive clustering and boundary recognition results, (c) the 3D image after robust filtering, (d) clustering and boundary recognition results.

5. Conclusion

The results of this study show that the combination of 3D robust smoothing filtering and adaptive clustering can effectively suppress Gaussian noise and outlier interference to obtain a stable and reasonable image segmentation results. The synthesis algorithm can be used not only in geophysical field, but also in other fields such as medicine, meteorology and communication.

6. Acknowledgement

This work was supported by Science Foundation of China University of Petroleum, Beijing (2462020YJRC010, 2462020YXZZ005) and the Key Laboratory of Geophysical Electromagnetic Probing Technologies of Ministry of Natural Resources (KLGEPT201908). The code is available by contacting the corresponding author.

References

[1] E. A. Guinness, R. E. Arvidson, J. W. Strebeck*et al.*, "Identification of a Precambrian rift through Missouri by digital image processing of geophysical and geological data," *J. Geophys. Res. Solid Earth.*, vol. 87, no. B10, pp. 8529-8545, Oct.1982.

[2] E. A. Guinness, R. E. Arvidson, C. E. Leff, *et al.*, "Digital image processing applied to analysis of geophysical and geochemical data for southern Missouri," *Econ Geol*, vol. 78, no. 4, pp. 654-663, Jul, 1983.

[3] D. Beamish, & D. K. Smythe, "Geophysical images of the deep crust: the Iapetus suture," *J GeolSoc London.*, vol. 143, no. 3, pp. 489-497, May, 1986.

[4] G. R. Pettifer, N. Djordjevic, D. Heislers, *et al.*, "Geophysical and image processing methods for detection of fireholes in brown coal, Latrobe Valley," *ExplorGeophys.*, vol. 20, no.2, pp. 153-158, 1989.

[5] D. Demanet, E. Pirard, F. Renardy, *et al.*, "Application and processing of geophysical images for mapping faults," *ComputGeosci*, vol. 27, no. 9, pp. 1031-1037, Nov. 2001.

[6] L. Zhang, T. Hao, J. Wu, *et al*. "Application of image enhancement techniques to potential field data". *Appl. Geophys.*, vol.2, no. 3, pp. 145-152, Sep, 2005.

[7] M. A. Simaan, Signal and image processing in exploration geophysics. *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp. 1-12, Feb. 2016.

[8] Z. Arifin, & A. Asano. "Image segmentation by histogram thresholding using hierarchical cluster analysis," Pattern Recognit. Lett., vol. 27, no. 13, pp. 1515-1521, Oct. 2006.

[9] S. A. Burney, & H. Tariq, "K-means cluster analysis for image segmentation," *Int. J. Comput. Appl.*, vol. 96, no. 4, pp. 2668-2677. Jun, 2014.

[10] J. Chen, H. Zheng, X. Lin, *et al.*, "A novel image segmentation method based on fast density clustering algorithm," *Eng. Appl. Artif. Intell.*, 73, pp. 92-110, Aug. 2018.

[11] B. Shi, L. Han, H. Yan, “Adaptive clustering algorithm based on KNN and density,” *Pattern Recognit. Lett.*, vol. 104, pp. 37-44. Mar, 2018.
[12] Dixit. “Adaptive kmeans clustering for color and gray image,” *MATLAB Central File Exchange.* Aug. 2020.
[13] V. Doré, & M. Cheriet, “Robust NL-means filter with optimal pixel-wise smoothing parameter for statistical image denoising,” *IEEE Trans. Signal Process,* vol. 57, no.5, pp. 1703-1716, Dec. 2008.
[14] Y. Wan, Q. Chen, Y. Yang, “Robust impulse noise variance estimation based on image histogram,” *IEEE Signal Process. Lett.*, 17(5), 485-488. Mar, 2010.
[15] M. P. Deisenroth, R. D. Turner, M. F. Huber, et al. “Robust filtering and smoothing with Gaussian processes,” *IEEE Trans. Automat. Contr.*, vol. 57, no.7, 1865-1871, Dec. 2011.
[16] Bhotto, M. Z. A., & Antoniou, A. (2011). Robust recursive least-squares adaptive-filtering algorithm for impulsive-noise environments. *IEEE Signal Process. Lett.*, 18(3), 185-188.
[17] Y. Chitour, R. Couillet and F. Pascal, "On the convergence of maronna’s m-estimators of scatter," *IEEE Signal Process. Lett.*, vol. 22, no. 6, pp. 709-712, Jun 2015.
[18] L. Chang and K. Li, "Unified form for the robust gaussian information filtering based on m-estimate," *IEEE Signal Process. Lett.*, vol. 24, no. 4, pp. 412-416, April 2017.
[19] P. J. Huber, Robust statistics (pp. 1248-1251). *Springer Berlin Heidelberg*. 2011.