

Research on adaptive trend surface filtering algorithm for multi-beam sounding data based on K-D tree

Fan Zhang^{1,*}, Xiaokai Li¹, and Shun Zhang¹

¹ College of Geomatics, Shandong University of Science and Technology, Qingdao 266000, China

Abstract. In the area with complex terrain changes, the traditional trend surface filtering has the problem that it is impossible to reasonably construct the terrain of the expression area. This paper proposes an adaptive trend surface filtering method based on K-D (K-Dimensional Tree) tree. Based on the K-D tree index, the algorithm divides the MBES (the multi-beam echo sounding system) data into several sub-blocks, and then analyzes each sub-block using trend surface filtering algorithm to more accurately reflect the real terrain. The experimental results show that the algorithm execution time in this case is about twice that of the traditional trend surface filtering in the case of millions of data volumes, and the execution efficiency is within a reasonable range. Compared with the traditional trend surface filtering algorithm, the algorithm has a higher fitting degree with the seabed terrain, and the depth difference distribution between the topographic point and the fitting plane is more concentrated. In addition, the proposed algorithm can effectively identify the outlier noise and the near-field noise in the case of ensuring the authenticity of the terrain, so it provides a useful reference for the denoising processing of MBES.

1 Introduction

In recent years, research on seabed information is a focus point in marine research [1]. With the development of acoustics, electronics, computers and other technologies, the shipborne multi-beam echo sounding system based on the echo detection principle has become a routine method for seabed topography detection. Due to the multi-beam sounding system's own noise, actual sea conditions, sound velocity profile and sonar parameters, there are a lot of noise in multi-beam point cloud data. Therefore, in order to obtain the real seabed topography and improve the accuracy of the seabed topographic map, the data collected by the multibeam sounding system must be refined.

Foreign scholars have done a lot of research on the problem of MBES data denoising. In the early days, mathematical statistics were mainly used to detect noise [2-3], such as Ware method [4] and Du method [5]. These methods are not only low in automation, but also the accuracy of the terrain and the accuracy of denoising are not ideal. In response to these problems, many experts have proposed a method combining mathematical statistics with spatial location for MBES data denoising. The most representative methods are threshold filtering [6], angle and gradient filtering [7-9], and local terrain fitting methods [10]. The research on domestic multi-beam data denoising theory is slightly later than abroad. Literature [11] proposed a trend surface filtering method, which uses a polynomial surface function to fit the seabed topography. This algorithm is simple to calculate and highly sensitive to

abnormal noises far from the terrain. However, the algorithm is not sensitive to the near-ground noise, the trend surface fitted by the algorithm is susceptible to noise and is not robust. In response to this problem, the literature [12] proposed the detection and elimination of abnormal sounding values based on the truncated least squares estimation. The method uses the high collapse point characteristic of truncated least squares estimation to perform local seabed trend surface fitting. However, the denoising effect of the algorithm and the algorithm execution efficiency are related to the size of the detection window. In practical applications, the window size needs to be manually and repeatedly adjusted. The literature [13] proposed an iterative trend surface fitting algorithm based on robust estimation. This algorithm has a high sensitivity to noise, but it needs to determine the weight of the massive point cloud and then iterate it, so it is difficult to ensure the operation efficiency. The literature [14] introduced the natural point influence domain and improved the trend surface filtering method. This method uses the autocorrelation coefficient as a condition for determining terrain consistency to perform local trend surface fitting. However, the algorithm needs to search for the natural neighbours of each sounding point, which makes the calculation more complicated, so it is not efficient for processing massive data.

Aiming at the above problems, this paper proposes a spatial adaptive trend surface filtering algorithm based on K-D tree index. The algorithm first establishes the K-D tree index of the 3D point cloud, and the K-nearest neighbour search is used to determine the neighbourhood information of the target point to block the bathymetric

* Corresponding author: fanzhang_cn@163.com

data. Then, the trend surface filtering analysis is performed on all the divided sub-blocks. In order to verify the effectiveness of the algorithm, design experiment compares the execution efficiency and denoising effect of the algorithm and traditional trend surface filtering algorithm.

2 Multi-beam sounding anomaly data elimination algorithm

2.1 Multi-beam outlier detection algorithm based on trend surface

The traditional trend surface filtering method adopts the principle of multiple regression to construct a reference trend that fit a mathematical surface to approximate the seabed topography. Then use this surface to simulate the spatial distribution and trend of terrain information. Finally, it is determined whether the measured value is an abnormal value according to the difference between the true measured value of the sounding point and the trend surface fitting value of the point.

The representation of the trend surface function is as shown in the following equation (1).

$$\hat{z}(x, y) = a_0 + a_1x + \dots + a_p y^k \quad (1)$$

Where, $P = k(k+3)/2$, a_0, a_1, \dots, a_p is a polynomial coefficient. Order:

$$X = \begin{bmatrix} 1 & X_{11} & \dots & X_{P1} \\ 1 & X_{12} & \dots & X_{P2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1N} & \dots & X_{PN} \end{bmatrix} Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} A = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{bmatrix}$$

The coefficient matrix can be obtained:

$$A = (X^T X)^{-1} X^T Z \quad (2)$$

The trend surface coefficient can be obtained from equation (2). Finally, statistically analyse the difference between the true measured value of the sounding point and the trend surface fitting value of the point. Points outside the standard range are defined as abnormal value to be eliminated. The trend surface filter is simple to calculate and the algorithm has better effect of eliminating outlier noise. It is suitable for flat sea area filtering where the terrain fluctuation is not obvious. The algorithm has poor detection ability for near-ground noise, and the denoising effect is affected by the filtering range, model order and detection principle. The main problem of the algorithm is how to construct the trend surface model reasonably. If the construction of the trend surface model is unreasonable, it will have a greater impact on the detection of outliers. Therefore, traditional trend surface filtering is only suitable for submarine terrain surveys with low accuracy requirements.

2.2 Adaptive trend surface filtering algorithm based on K-D tree

Traditional trend surface filtering is highly efficient and can effectively identify and eliminate significant abnormal noise. However, in the trend analysis process, the selection of the order of the polynomial is directly related to the effect of the trend surface filtering. When the order of the polynomial is too high, not only the amount of calculation is greatly increased, but also the abnormal value may be retained. When the order of the polynomial is too low, the fitted surface may deviate from the real terrain, resulting in loss of effective information, and it is difficult to effectively eliminate noise. In addition, when the terrain changes in the fitted area are complicated, it is difficult to construct a suitable trend surface representation area topography. In recent years, the research on trend surface filtering has become a hot topic in MBES data post-processing theory research. Therefore, the targeted improvement of traditional trend surface filtering algorithm has important significance and practical application value for MBES data post-processing work.

This paper proposes a spatial adaptive trend surface filtering algorithm based on K-D tree index. The basic idea is to divide the sounding data into several sub-blocks and perform trend-surface filtering analysis on each sub-block separately. In this way, the trend surface fitted by each sub-block can better reflect the real terrain compared with the overall trend of the traditional trend surface filtering. The algorithm is also more capable of retaining valid information and detecting outliers. The algorithm first establishes the K-D tree index of the 3D point cloud, and then determines the neighbourhood information of the target point. While using the neighbourhood information for trend surface fitting, the difference between the true measured value of the sounding point in the neighbourhood and the trend surface fitting value of the point is statistically calculated. When the difference between the depth value of the sounding point and the trend surface fitting value of the point is outside the standard range, the point is defined as abnormal value and removed. After completing the above steps, go to the next target point and perform trend surface detection again until all data in the point cloud set passes the trend surface detection. Traditional trend surface filtering uses a fitting trend to represent the spatial distribution of terrain features. However, for each sub-block of spatial adaptive trend surface filtering, not only the trend surface parameters need to be adjusted, but also the order of the polynomial of the trend surface needs to be selected. The fitness index R^2 of the trend surface and the actual surface is usually used as an indicator to measure the degree of fit between the trend surface and the real terrain.

$$R^2 = 1 - \frac{\sum_{i=1}^n (z_i - \hat{z}_i)^2}{\sum_{i=1}^n (z_i - \bar{z})^2} \quad (3)$$

Among them, z_i is the measured true value of the sounding point, \hat{z}_i is the trend value of the sounding point, and \bar{z} is the average depth value. The higher the value of R^2 , the better the fit between the fitting trend and the real terrain. The following describes the implementation steps of the algorithm and the implementation process of the algorithm.

2.2.1 Algorithm steps

The algorithm steps are as follows:

1) This algorithm needs to input the point cloud *cloudPoint* first and create the KD index tree. Then walk through every point *cloudPoint* in *searchPoint*, doing steps 2)-9). After traversal, turn 10).

2) Check if *searchPoint* has been searched; check if *searchPoint* is an invalid point; if the above two meet one of them, jump out of this cycle, turn 1), otherwise, turn 3).

3) Setting a search radius r and performing a k-neighbour search on *searchPoint*, the searched points are recorded as K_Num , and the searched point and the index number corresponding to *searchPoint* are saved in the vector container *pointVector*.

4) Save the index number of the valid data point in *pointVector* and the tag value 0 to the hash table 1 to determine whether the current point is searched.

5) Set threshold λ_1 , if $K_Num \leq \lambda_1$, it is determined that *pointVector* corresponds to the point in *cloudPoint* is an invalid point, marked as invalid point and stores the index number of the point and the distance value 0 to the hash table 2, turn 1); otherwise, turn 6);

6) Determine the order of the trend surface to be fitted, and perform a trend surface fitting on the points in *pointVector* and the trend surface is fitted as A .

7) Calculating a difference between a depth value of a point corresponding to the index number in *pointVector* and a trend surface fitting value of the point, and saving the index number and the corresponding difference value.

8) Calculate the mean and medium error where the difference is not 0, denoted as u and σ respectively

9) According to the prese threshold λ , the difference is traversed. If the difference is between $[u - \lambda \cdot \sigma \quad \mu + \lambda \cdot \sigma]$ and the point is not marked, the point in the index corresponding to the index number *cloudPoint* is marked as a topographic point, and the index number of the data point and the marker value 0 are saved to Hash table 3.

10) Save topographic points and noise data separately to the PCD format file.

2.2.2 Algorithm Flow

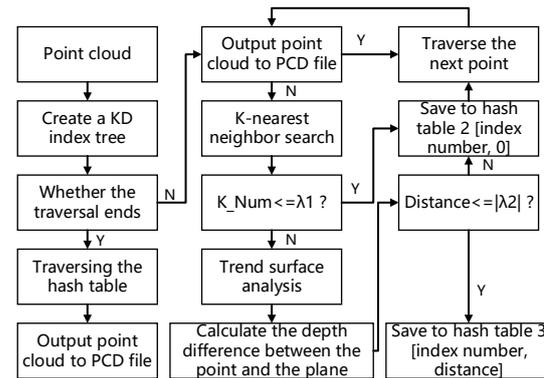


Fig. 1. Adaptive trend surface filtering algorithm flow

3 Experimental design and results analysis

The experimental data is derived from the measured data of a coastal zone in the Yangtze River estuary in 2018. The field investigation tasks include 500km multi-beam seabed topography and full coverage measurement survey, backscatter data processing results and sediment classification results in key areas, and near-shore underwater points in a detailed investigation area of the East China Sea. The experimental data is the measured point cloud collected by the R2Sonic2024 multi-beam system, and the measurement range is about 3km². The experimental hardware environment is CPU: 3.5GHz, Intel(R) Core (TM) i3-4150, memory: 8GB. The operating system is: Ubuntu 16.04.

In the experiment, the original data is firstly corrected by the CARIS HIPS software, the sound speed is corrected, and the data is fused. Convert the data format of the processing result, and finally read the converted data in the program for filtering. The above two filtering algorithms are used to process the same set of point cloud data, and the two algorithms are compared and analysed from the point cloud denoising efficiency and the point cloud denoising effect.

3.1 Formatting the text

The above two filtering algorithms are used to process the same set of point cloud data, and the traditional trend surface filtering is named as algorithm one, and the adaptive trend surface filtering algorithm based on K-D tree index is named algorithm 2. The Parameter of algorithm 1 is set to $k = 2.0$, and the Parameter of algorithm 2 are set to $K = 1000$, $k = 2.0$. The statistical results of each algorithm are shown in Table 1.

The execution efficiency of adaptive trend surface filtering is related to the size and number of filtering sub-blocks. When using the formula (Algorithm 2 execution time - Algorithm 1 execution time) / Algorithm 1 execution time to calculate millions of levels of data in a table, the result is 115.0%, 91.3%, 98.8%, 103.5%, and 96.5%. The experimental results show that the algorithm

runs at twice the time of the traditional trend filter in the case of millions of data.

Table 1. Filter algorithm execution time statistics

Points	Algorithm 1 /s	Algorithm 2 /s
5.5×10^5	2.843	5.146
1.0×10^6	5.157	11.086
2.2×10^6	11.462	21.208
4.0×10^6	20.817	40.123
6.3×10^6	31.325	63.749
8.2×10^6	41.357	81.287
1.0×10^7	56.114	103.801
1.2×10^7	71.524	125.660

3.2 Algorithm denoising effect analysis

This paper selects a representative terrain and demonstrates the denoising effect from both the point cloud denoising effect and the point cloud statistical analysis. The following is a visual display of the denoising effect of the point cloud.

3.2.1 Formatting the title Point cloud denoising effect

(1) Traditional trend surface filtering

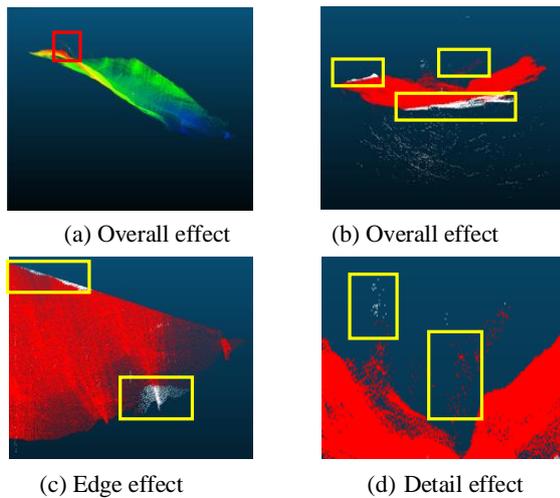


Fig. 2. Regional point cloud effect

Figure 2 is a point cloud effect diagram processed by the trend surface filtering algorithm. It can be seen from Fig. 2(a) and Fig. 2(b) that the trend surface filtering can eliminate most of the outlier noise, but there is a large amount of real terrain information loss in the boundary area. It can be found from Fig. 2(c) that some boundary point clouds are deleted by mistake. From Figure 2(d), it can be found that there is a clear boundary between the topographic point and the noise data judged by the algorithm. The trend surface filtering cannot effectively identify the cluster noise, and most of the cluster noise is mistaken for the terrain by the algorithm. So that it can be retained.

(2) Adaptive trend surface filtering algorithm based on K-D tree index

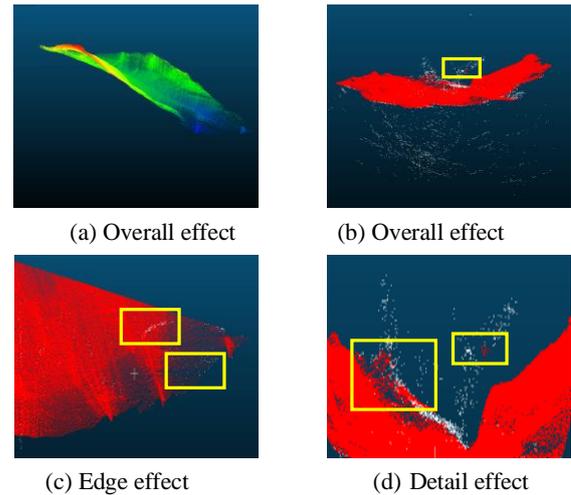


Fig. 3. Regional point cloud effect

Figure 3 is a point cloud effect diagram processed by the Adaptive trend surface filtering algorithm based on K-D tree index. It can be seen from Figure 3(a) and Figure 3(b) that adaptive trend surface filtering can identify and exclude a large amount of noise. From Figure 3(c), it can be found that the boundary retention is improved compared to the traditional trend surface filtering, and only a small number of boundary points are deleted by mistake. From Figure 3(d), it can be found that compared with the traditional trend surface filtering, the adaptive trend surface filter has stronger ability to identify noise, and there is no obvious boundary between the terrain point and the noise data.

3.2.2 Point cloud statistical analysis

The first type of error refers to the ratio at which the algorithm mistakes the noise point for terrain. The second type of error refers to the ratio of the algorithm to the topographical point as a noise point. Using the results of human interaction processing as a standard, the terrain data selected by the two algorithms is analysed. The results are shown in Table 2.

Table 2. Filter algorithm error rate statistics

	Algorithm 1 /s	Algorithm 2 /s
Maximum	190.973	190.973
Minimum value	166.472	169.300
Average value	178.038	178.042
One type of error rate	14.651%	6.767%
Second type error rate	0.554%	0.276%

The first type of error rate of the traditional trend surface filtering is the highest, indicating that the algorithm has poor ability to recognize noise. The first type of error rate of adaptive trend surface filtering is lower than that of traditional trend surface filtering, and the algorithm has stronger ability to recognize noise. The second type of error rate of traditional trend surface filtering is the highest, indicating that the ability of the algorithm to retain the accuracy of terrain authenticity is poor. The second type of error rate of for adaptive trend

surface filtering is lower than that of traditional trend surface filtering, and indicating that the algorithm is more capable of retaining real terrain.

4 Conclusion

This paper aims at the problem that the traditional trend surface filtering algorithm cannot reasonably construct the expression of regional reference trend in a complex environment. Based on the traditional trend surface filtering algorithm and k-d tree index, an adaptive trend surface filtering algorithm based on k-d tree is proposed. The experimental results show that the proposed algorithm has good denoising effect, and compared with the traditional trend surface filtering algorithm. The ability of the algorithm to preserve real terrain is greatly improved. However, Algorithm execution time is twice that of traditional trend surface the execution efficiency of this algorithm is twice that of the traditional trend surface. Therefore, this paper will study how to improve the efficiency of execution.

References

1. Wu. Ziyin. Linear sand ridges on the outer shelf of the East China Sea. *Chinese Science Bulletin*, **21**, 2517-2518, (2005).
2. Ladner. R W, Elmore. P, Perkins. A L. Automated cleaning and uncertainty attribution of archival bathymetry based on a priori knowledge. *Marine Geophysical Research*, **38(3)**, 291-301, (2017).
3. Shaw. S, Arnold. J. Automated error detection in multibeam bathymetry data, *Oceans IEEE*, **2**, II/89-II/94, (1993).
4. Ware. C. A system for cleaning high volume bathymetry, *International Hydrographic Review*, **2**, 77-94, (1992).
5. Du. Z, Wells. D, Mayer L. An approach to automatic detection of outliers in multibeam echo sounding data. *Hydrographic Journal*, **79(1)**, 19-25, (1996).
6. Rezvani. M H, Sabbagh. A, Ardalan. A A. Robust Automatic Reduction of Multibeam Bathymetric Data Based on M-Estimators, *Marine Geodesy*, **38:4**, 327-344, (2015).
7. Bourillet. J F, Edy. C, Rambert F. Swath mapping system processing: Bathymetry and cartography. *Marine Geophysical Researches*, **18(2-4)**, 487-506, (1996).
8. Caress. D W, Chayes. D N. Improved processing of Hydrosweep DS multibeam data on the R/V Maurice Ewing. *Marine Geophysical Researches*, **18(6)**, 631-650, (1996).
9. Pereda García, Raúl, et al. Model for the Processing and Estimation of Dual Frequency Echo Sounder Observations in Detailed Bathymetries. *Marine Geodesy*, **39(3-4)**, 305-320, (2016).
10. Debes. N. Multibeam Echosounder Data Cleaning Through an Adaptive Surface-based Approach, *Computers & Geosciences*, **46**, 330-339, (2007).
11. Zhu. Q, Li. D. Error analysis and processing of multibeam soundings, *Journal of Wuhan Technical University of Surveying and Mapping*, **1**, 3-6+48, (1998).
12. Lu D, Li H Wei Y. A method of multi-beam bathymetry outliers elimination base on trimmed least squares estimation, *Journal of Geodesy and Geodynamics*, **32(1)**, 89-93, (2012).
13. Zhang. Z, Bao. J, Xiao. F. Application of robust estimation for detecting outliers of multibeam data, *Journal of Liaoning Technical University (Natural Science)*, **35(7)**, 755-758, (2016).
14. Zhang Z, Peng R, Huang W. An improved algorithm beam bathymetric data considering the natural neighboring points influence field, *Acta Geodaetica et Cartographica Sinica*, **1**, 35-47, (2018).