Research and Implementation of Modeling Grid DEM Based on Discrete Data

Jing Wang, Kai Xia, Xuegong Chen

Research and Implementation of modeling Grid DEM based on discrete data

Jing Wang[1,a]， Kai Xia[1,b], Xuegong Chen[1,c]

[1] Software School, Central South University, Changsha 430070, China
[a]weyure@126.com, [b]xiakai05@163.com, [c]cxgcsu@csu.edu.cn

Abstract. Interpolation method is used to construct regular grid DEM (Digital Elevation Model) based on discrete data. The key to DEM interpolation algorithm lies in the efficiency of data search around the points to be operated and the precision of DEM interpolation by selecting an appropriate interpolation function model. At present, the point by point interpolation is very common. Based on the comparison and analysis of characteristics and application scope of various DEM interpolation algorithms, this paper proposed a new interpolation model which combines distance-weighted average method and least squares fitting method. Results show that proposed method can improve search speed and interpolation accuracy remarkably and has good application value for large area.

Keywords: DEM interpolation, digital elevation model, distance-weighted average method, least squares fitting method

1   Introduction

With the rapid development of GIS (Geographic Information System), DEM (Digital Elevation Model), as a digital description and simulation of the surface topography, has become an important component of spatial data infrastructure and "Digital Earth". The accuracy of DEM has a significant impact on terrain visualization, GIS analysis and credibility of the decision-making [1, 2]. It is necessary to improve the accuracy of modeling DEM. Most of original data are discrete data. The key issue of DEM is how to construct regular grid DEM by interpolating with discrete data. Therefore, research on improving the DEM accuracy is of great significance.

To construct Grid DEM, first we divide the research region into two-dimensional grid to form spatial structure of grid that covers the entire area. Then discrete points surrounding the interpolated grid points are used to calculate its elevation values, we can obtain the grid data of this region with a certain format. Interpolation is the core issue of modeling DEM, which involved in DEM in all aspects of production, including quality control, accuracy assessment and analysis applications. According to the distributed range of interpolation points, DEM data interpolation algorithm generally can be divided into overall interpolation, block interpolation and point by point interpolation. The point by point interpolation is widely used because its calculation method has many advantages, such as simple, flexible block, high precision and so on. The point by point interpolation can be divided into the distance-

weighted average method, moving surface fitting method, least squares fitting method, etc. Every interpolation method has some drawbacks. For example, the distance-weighted average method is less precise and poor in smoothness, and only suitable for points which are well-distributed and sampled from even terrain; Moving surface fitting method is of high precision and flexible in calculation, but imposes high requirements on the sampling point; Least squares fitting method is suitable for most of terrain but time consuming [3-8]. Different interpolation algorithms have their own advantages, disadvantages and application scope, so we should select the appropriate interpolation algorithm according to a certain problem.

Simply using one method of interpolation to construct DEM in entire region have limitations, such as high requirements on the sampling point, time consuming and so on, which will affect the applicability and accuracy of model. People put forward some combination method, for example, integration of moving surface fitting method with distance-weighted average method, distance-weighted average method with quadratic surface method. Both of these methods only suitable for points which are well- distributed and sampled from even terrain. The method that combines moving surface fitting method with least squares fitting method does not consider moving surface method has a high requirement on accuracy of sampling data [9-13].

In response to above problems, taking all conditions into account and after analysis of the application of various interpolation methods, this paper presents a method that integrates the distance-weighted average method with the least- squares fitting to solve problems arising in DEM modeling and improve accuracy and efficiency of the interpolation method.

The following of the paper is organized as follows. Section 2 presents the proposed approach. Section 3 describes an experiment and shows the analysis. Section 4 concludes the paper.

2   Data and Methods

2.1   Data

These elevation data used in this paper are sampled from three areas, such as alpine areas, glacier areas and even areas. Each area contains more than 50000 sampling points. Using the distance-weighted average method, least squares method, moving surface fitting method, moving surface fitting method and the proposed combination method to construct DEM based on discrete data, we analyze and evaluate the accuracy of the model.

Set aside 100 sampling points as the checkpoints that do not participate in DEM modeling. Using different methods to model DEM, comparing the checkpoints original elevation values and calculated elevation values through four interpolation algorithm to calculate max error and RMSE (root mean square error), then we analyze and evaluate the results.

2.2   Distance- weighted average method

To use the distance-weighted average method, we need to select these discrete points nearest to the interpolation point. The easiest way is to directly calculate the distance between interpolation points and the discrete points, then choose discrete points according to distance and interpolate, but this interpolation method need large time and space to compute, which has a bad impact on interpolation speed [10]. So we choose certain search areas according to requirement of mathematical model for the interpolation number and adjust the search scope.

The distance-weighted average method is as follows:

$$h = \frac{\sum_{k=1}^{m} w_k h_k}{\sum_{k=1}^{m} w_k}$$

(1)

h_k is the point to be interpolated, w_k is the weight of discrete point k, and h_k is elevation values of sampling point k.

Currently weight calculated in the following ways:

$$P_i = d_{ki}^{-u}$$

(2)

$$P_i = \left(\frac{R - d_{ki}}{d_{ki}} \right)^2$$

(3)

$$P_i = e^{\left(\frac{d_{ki}}{m} \right)^2}$$

(4)

R is a search radius; m is a constant, P_i is the weight of discrete points. In this paper, we choose the first method, and set u = 2.

A problem arises when there are few reference points. When the reference points are very far from each other or distributed unevenly, if we still use these reference points that has little effect on the interpolated points to interpolate, it will cause disturbance of interpolation points instead of improve the accuracy of the interpolation points, which make the interpolation results not ideal.

To improve the efficiency of distance-weighted average algorithm, this paper adopts the idea of gradually expanding the search area and restricted orientation search. The idea is as follows:

First, classify grid in accordance with the actual needs.

Second, determine the number of discrete points needed to be calculated.

Third, search discrete points around the interpolated grid points with the restricted orientation. If there are enough discrete points in each quadrant, use the distance-weighted average method for interpolation. If not, expand search area gradually. When search area reaches threshold, then stop.

2.2 Least squares fitting method

When the reference points are very far from the interpolation points or distributed unevenly, distance-weighted average method has great error. Least squares fitting method, which combines the advantage of distance-weighted average method and trend surface method, is suitable for most of terrain, and can effectively eliminate noise data [12].

Weighted least squares fitting method assigns all discrete points that have influence on interpolation points a certain influence coefficients, and assumes the coefficients is consistent with a quadratic polynomial as follows:

$$p(x,y) = c_1 + c_2 x + c_3 y + c_4 xy + c_5 x^2 + c_6 y^2 \tag{5}$$

$$Q = \sum_{i=2}^{n} [p(x_i, y_i) - z_i] w [(x_i - a)^2 + (y_i - b)^2] \tag{6}$$

When we use least squares fitting method for interpolation, these points that are nearer to the interpolation point are allocated a higher weight than these points which are far away. As the number of discrete points the least squares fitting method requires is more than 6，we need to find coefficients ci (i= 1, 2, $\cdots$, 6) to make (6) minimum. $w[(x_i - a)^2 + (y_i - b)^2]$ is weight and is a function of the distance. Take $w[(x_i - a)^2 + (y_i - b)^2] = \dfrac{1}{[(x_i - y_i)^2 + (y_i - b)^2]^2}$, when （a, b）close to (x_i, y_i), it has a large value and a small value while （a, b）is far from (x_i, y_i). To get minimal Q, according to the principle of least squares $\dfrac{\partial Q}{\partial c_i} = 0$, so we can get the following equation:

$$\frac{\partial Q}{\partial c_1} = \sum_{i=1}^{N} 2(P(x_i, y_i) - z_i) w = 0 \tag{7}$$

$$\frac{\partial Q}{\partial c_2} = \sum_{i=1}^{N} 2(P(x_i, y_i) - z_i) w x_i = 0 \tag{8}$$

$$\frac{\partial Q}{\partial c_3} = \sum_{i=1}^{N} 2\left(P(x_i, y_i) - z_i\right) w y_i = 0 \tag{9}$$

$$\frac{\partial Q}{\partial c_4} = \sum_{i=1}^{N} 2\left(P(x_i, y_i) - z_i\right) w x_i y_i = 0 \tag{10}$$

$$\frac{\partial Q}{\partial c_5} = \sum_{i=1}^{N} 2\left(P(x_i, y_i) - z_i\right) w x_i^2 = 0 \tag{11}$$

$$\frac{\partial Q}{\partial c_6} = \sum_{i=1}^{N} 2\left(P(x_i, y_i) - z_i\right) w y_i^2 = 0 \tag{12}$$

The equations from (7) to (12) can be solved to obtain coefficients of p (x_i, y_i). Use these coefficients in equation (5); we can obtain value of grid points. Solving these equations needs large time. When there are large discrete points, there is a sharp decline in speed. When the reference points are very far from the interpolation points or distributed unevenly, least squares fitting method has better accuracy than distance-weighted average method. So we combine least squares fitting method and distance-weighted average method to improve the accuracy and efficiency of interpolation.

2.2   Proposed method

When the reference points are very far from each other or distributed unevenly, the result of using weighted average algorithm would not be ideal. This paper brings up the idea of gradually expanding the search area and restricted orientation search. As the number of discrete data points the distance-weighted average method requires is from 4 to 8, we adopt four quadrants search method, so there should be at least one point in each quadrant. The least squares fitting method has no special requirements on the distribution of discrete points, but it needs at least six discrete points. When the number of discrete points is over six and the distribution does not satisfy the requirements of distance-weighted average method, we adopt least squares fitting method.

In addition, in order to accelerate the computing speed and improve the efficiency of modeling, we judge distance between point to be interpolated and discrete point firstly. For example, if the nearest distance between point to be interpolated and discrete point is less than 0.5 meter, then we use value of the elevation of discrete point directly and don't do next step, if not, use the proposed algorithm for interpolation. Algorithm is as follows:

First, for point P to be interpolated, search discrete points in grid with number 1 around the point to be operated with the restricted orientation searching. If each of four quadrants has more than one discrete point, use distance-weighted average method for interpolation；

Second，if not, expand search area with number 2. Check if there is more than one discrete point in each quadrant, if yes; use distance-weighted average method for interpolation；

Third, if discrete points in one or more quadrants are less than one, but a total of discrete points are more than six, use least squares fitting method to interpolate.

Fourth, if discrete points are less than 6, expand the search area with number 3 and use least squares fitting method to interpolate if the requirement is satisfied ;

Sixth, if not, don't do interpolation and process the next grid point.
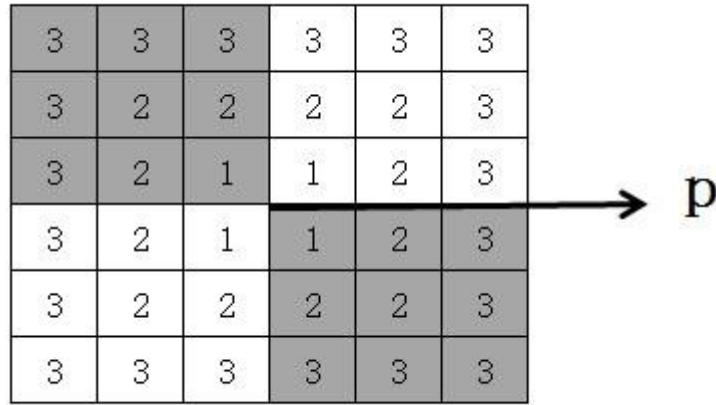

Fig. 1. Grid model searching

3 Experiments and Discussion

3.1 Experiments

Using distance-weighted average method, least squares fitting method, moving surface fitting method and proposed method to model DEM. Set aside 100 sampling points as the checkpoints that do not participate in DEM modeling, if the check point distribute unevenly, we may get an inaccurate evaluation of interpolations, in this paper we make the check point distribute evenly in order to obtain an accurate evaluation. We compare the original elevation values and calculated elevation values through four interpolation algorithms to calculate RMSE and max error. In order to directly display the results, we randomly selected 10 sampling points from 100 sampling points to compare. Fig. 2, Fig. 3 and Fig. 4 show absolute value of error in

alpine areas, glacier areas and even areas respectively. Table 1, table 2 and table 3 show comparison of various interpolation precision in alpine areas glacier areas and even areas respectively. The results are as following：
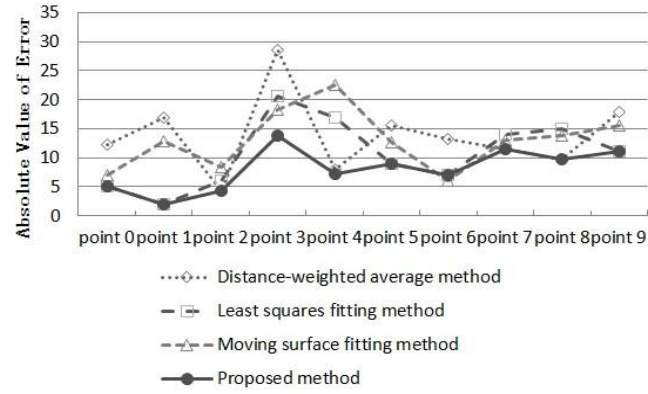

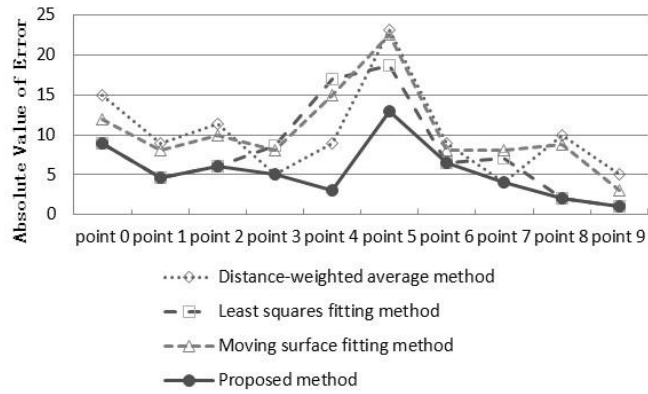**Fig. 2.** Absolute value of error in alpine areas

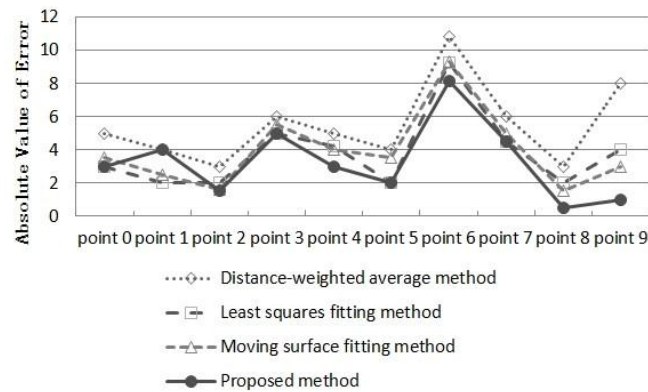**Fig. 3.** Absolute value of error in glacier areas

Fig. 4. Absolute value of error in even areas

Table 1. Comparison of various interpolation precision in alpine areas (unit: meter)

Interpolation method	Distance-weighted average method	Least squares fitting method	Moving surface fitting method	Proposed method
RMSE	15.27	12.06	14.11	8.48
Max error	28.67	20.64	22.53	13.77

Table 2. Comparison of various interpolation precision in glacier areas (unit: meter)

Interpolation method	Distance-weighted average method	Least squares fitting method	Moving surface fitting method	Proposed method
RMSE	11.36	9.91	12.24	6.54
Max error	23.13	18.71	27.15	12.92

Table 3. Comparison of various interpolation precision in even areas (unit: meter)

Interpolation method	Distance-weighted average method	Least squares fitting method	Moving surface fitting method	Proposed method
RMSE	5.78	4.46	4.51	3.37
Max error	10.81	9.25	9.33	8.16

3.2 Discussion

From these tables, we can sum up:

The proposed method has higher accuracy than distance-weighted average method and least squares fitting method:

In alpine areas, RMSE of using proposed method is 8.48, RMSE of using least squares fitting method is 12.06, RMSE of using distance-weighted average method is 15.27, and RMSE of using moving surface fitting method is 14.11, an decrease of 3.58, 6.79, 5.63 respectively.

In glacier area, RMSE of using proposed method is 6.54, RMSE of using least squares fitting method is 9.91, RMSE of using distance-weighted average method is 11.36, and RMSE of using moving surface fitting method is 12.24, an decrease of 3.37, 4.82, 5.7 respectively.

In even areas, RMSE of using proposed method is 3.37, RMSE of using least squares fitting method is 4.46, RMSE of using distance-weighted average method is 5.78, and RMSE of using moving surface fitting method is 4.51, an decrease of 1.09, 2.41, 1.14 respectively.

This result amply demonstrated that the proposed method has significantly improved accuracy in various areas, especially in alpine areas. This proved that the proposed method has a great of advantage in meeting actual demand and particularly suitable for the modeling of uneven terrain of the area.

The distance-weighted average method and moving surface fitting method has weak accuracy in alpine areas and glacier areas, better accuracy in even areas. The least squares fitting method has higher accuracy than distance-weighted average method in each kind of areas, especially in undulating areas. But in even area,

accuracy is only a few higher. Taking the computing time into account, the least squares method is suitable for terrain that changes significantly.

4 Conclusions

Using distance-weighted average method, least squares fitting method, moving surface fitting method and proposed method to model DEM, according to the experimental results, after analysis, we can get the following conclusions.

The proposed method which combines distance-weighted average method and least squares fitting method has adopt both advantages and improved accuracy in each kind of terrain; it is suitable for the modeling of large areas. The experiment has demonstrated that the proposed method has improved efficiency of interpolation and prevented generating unwanted interpolation points.

The thinking of distance-weighted average method takes large ground gradient as a continuous process. But when the terrain changes significantly, such as ridges or rifts, the accuracy is not particularly ideal. So this method is suitable for even areas.

The least squares fitting method has high accuracy and time consuming, so it is suitable for terrain changes significantly.

Moving surface fitting method is suitable for points which are well- distributed and sampled from even terrain. When the reference points are very far from each other or distributed unevenly, this method would not be ideal, because it has a high requirement on accuracy of sampling data.

It is difficult to find a model that fully fit the complex topography of the surface. Therefore, in practical applications, how to make better use of space to fit the data and actual ground is an important direction for future research.

References

1. Li Zhilin, Zhu Qing. Digital Elevation Model (Second Edition) [M]. Wuhan: Wuhan University Press, 2003, 6-7
2. Roger H Twito, Ronald W Mifflin, Robert J McGaughey. The MAP Program：Building the Digital Terrain Model [E]. US: Pacific Northwest Research Station，2000
3. Tang Guoan, Li Fayuan, Liu Xuejun. Digital Elevation Model (Second Edition) [M]. Wuhan: Science Press，2005, 76-85
4. Julie Charleux Demargne, Christian Puech. Quality assessment for drainage networks and watershed boundaries extraction from a digital elevation model (DEM) [J]. Proceedings of the 8th ACM international symposium on Advances in geographic information systems，2011, 5: 89~94
5. Fernando J. Aguilar, Francisco Agüera, Manuel A. Aguilar et al. Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy [J]. Photogrammetric Engineering & Remote Sensing, 2005, 7:805-816
6. GL Heritage, DJ Milan, ARG Large.IC. Fuller influence of survey strategy and interpolation model on DEM quality [J]. Geomorphology, 2009, 112(3):333-344

7. Lazzaro, D, L.B. Montefusco. Radial basis functions for the multivariate interpolation of large scattered data sets [J]. Journal of Computational and Applied Mathematics, 2002, 140(1–2):521–536.

8. Li, J, and C.S. Chen. A simple efficient algorithm for interpolation between different grids in both 2D and 3D. Mathematics and Computers in Simulation, 2002 58(2):125–132

9. Pan Pan, Wang Guangxia, Zhang Hua. Study and practice of DEM generation [J].Journal of Zhengzhou Institute of Surveying and Mapping, 2007, 24(1):57-60

10. Li Yin, Yang Wunian, Yang Ronghao et al. Improvement of DEM interpolation algorithms based on moving surface fitting and distance-weighted [J]. Map, 2010, 33(4): 168-171

11. Wang Jianming, Yang Xiaoqin, Yao Jianhai. Integration of the two methods to interpolate grid DEM [J]. Journal of Taiyuan university of technology, 2008, 5(39):63-69

12. Hu Jie. The study of DEM modeling method based on regular grid [J]. Geomatics and spatial information technology, 2012, 35(11):138-140

13. Zhang Jinming, Guo Liping, Zhang Xiaodan. Effects of interpolation parameters in inverse distance weighted method on DEM accuracy [J]. Journal of Geomatics Science and Technology, 2012, 1:51-56.