International Journal of Multimedia and Ubiquitous Engineering
Vol.9, No.5 (2014), pp.185-194
http://dx.doi.org/10.14257/ijmue.2014.9.5.18

A Simplification Method of Terrain Modeling Based on Spatial-autocorrelation

Yunhua Gu, Xiaoyan Wu and Jin Wang

Jiangsu Engineering Center of Network Monitoring, School of Computer & Software Nanjing University of Information Science and Technology, Nanjing, 210044, China

Abstract

To solve the problem of three-dimensional terrain simplification in large-scale which was based on digital elevation model (DEM) data, a terrain simplified method based on the theory of spatial autocorrelation was proposed. The slope, a key factor for expressing the terrain features, has been taken into account in this method. According to principle of regional similarities in geomorphology is related to DEM data correlation, a cluster analysis was performed on the terrain data to get the thresholds distinguish terrain features. On the basis, the gradient-based distance-weighted method was adopted to fit elevation values of the center point and generated a new terrain mesh. The results of experiments show that the method to some extent reduces the size of the data, while maintaining good terrain features and curvature characteristic.

Keywords: *spatial-autocorrelation, terrain simplification, Digital Elevation Model, gradient, distance-weighted*

1. Introduction

Digital elevation model (DEM) as a basis of DGM, the data structure of which is beneficial to implement the terrain analysis, is a digital description of terrain and locates the elements of terrains by the two-dimensional geospatial [1]. However, the overlarge cell of the DEM grid will has a negatively influencing for the expression of terrain characteristics and has a data redundancy. In the construction of the massive three-dimensional terrain, because of higher resolution of the elevation data grid and the wider geographical scope, these factors will result in a tremendous amount of data which will be used for modeling three-dimensional terrain. It will require a better terrain-modeling technique and power computing ability for amount of computing, in this case. For example, a large terrain, latitude and longitude difference of 5°×5°, 555 Km×555Km range, the sampling points of DEM data is huge, and if these points are not simplified, the performance of the graphic will be difficult to meet the practical needs. So, the simplification for the terrain is the basis of constructing a three-dimensional GIS, especially the data of which is based on DEM [2].

There are numerous studies dealing with the simplification for the large terrain modeling. Literature [3] proposed new algorithms for simplifying terrain surfaces designed specifically for a new measure of quality based on preserving inter-point distances; Literature [4] proposed a new 3D terrain mesh simplification method, which is based on analysis of clustering and Octree to organize the streamlining processes, achieving the model transition from simple to complex. Literature [5], based on the Garland algorithm, adopts the algorithm which is based on the feature degree of a vertex to simplify a mesh model, to solve the problems of losing some important shape features at low levels of the simplified model, such

ISSN: 1975-0080 IJMUE

as valley and channel. Literature [6] propose and discuss a new Lepp-surface method able to produce a small triangular approximation of huge sets of terrain grid data by using a two-goal strategy that assures both small approximation error and well-shaped 3D triangles. Literature [7] manually selects some important terrain features for the purpose of optimizing the geometric model, according to the actual needs. Literature [8] introduces a real-time LOD method which divides the terrain in a set of nested regular grids centered about the viewer.

As can be seen from the above, there still exist some problems. In terms of display efficiency, the original data has been reorganized and adopt hierarchical scheduling algorithm to improve the performance of displaying, but without reducing the overall amount of data; In the terms of terrain's minutiae, features of terrain is described using geometric model which is within the permitted range of error. If characteristic values are determined by a predetermined way in the processing, it will be less flexible in practical applications; If the characteristic points are determined by surface curvature, the results will be more accurate, but increasing the computational complexity through the modeling process.

To solve the problem of three-dimensional terrain simplification in large-scale which was based on digital elevation model (DEM) data, a terrain simplified method based on the theory of spatial autocorrelation was proposed in this paper. Landform is a collection of the natural topography fluctuation state, according to the geography description. The slope is an important means to describe the terrain, and meanwhile the things in a geographical space meet the autocorrelation. Hence, this paper firstly has a clustering analysis for elevation data because of the spatial autocorrelation property, so as to find out the terrain characteristic value; secondly, to compute the elevation values of the center point by the gradient-based distance-weighted algorithm and reconstruct the original data mesh. The experiment shows that the method can reduce the scale of data and maintain good terrain features; meanwhile, the curvature changes a little compared to the original result.

The remainder of this paper is organized as follows: Section 2 discusses the digital elevation model with spatial autocorrelation theory in our model. Section 3 describes the method of clustering analysis for topographic. Section 4 gives the description of the gradient-based distance-weighted algorithm that we proposed. Section 5 gives the processes of terrain simplification. Section 6 discusses the experiments and analysis. Section 7 concludes our work and future work.

2. Digital Elevation Model with Spatial autocorrelation Theory

In the following section, we will describe the digital elevation model with spatial autocorrelation theory. Digital Elevation Model can be described as follows: {X, Y and H}, X: longitude, Y: latitude. X and Y determine the location of points at the two-dimensional terrain surface, and H is the elevation value. Spatial autocorrelation theory explicates the relationships of adjacent-space elements between properties and position. The theory measures the degree of similarity or dissimilarity of the spatial attribute values, and reflects some certain geographical phenomenon in an area. Elevation value as a spatial property also can be measured by the theory above. The autocorrelation characteristics of local-adjacent positions affect the degree of redundancy of elevation data, the closer the relationship, and the more redundant. If the area has fewer minutiae, the redundancy will be more in elevation data than an area having more minutiae. On the basis of discussing above, we can use this feature to simplify the original grid data.

3. Clustering analysis for topographic

Landscape is the natural undulating state of the ground, and the slope is a key factor for describing the surface of terrain. We can use the ratio of elevation difference to horizontal distance between two points. So, the type of land, such as mountains and plains, can be easily divided on the basis of this terrain factor. However, the time consuming will be huge for calculating the slope value in terrain constructing. For this reason, we use the K-means clustering analysis to classify the elevation value.

Partitioned clustering techniques can create a level partitioning of data points. There are a number of partitioned techniques. K-means is based on the idea that a center point can represent a cluster. Especially, for K-means we use the notion of a centroid, which is the mean or median point of a group of points. K-means method is based on similarity measure between samples and is indirect clustering method. The basic K-means clustering technique is presented below.

Step1. Select K points as the initial centroids.

Step2. Assign all points to the closet centroid.

Step3. Re-compute the centroid of each cluster.

Step4. Repeat steps2 and 3 until the centroids don't change.

The elevation data is logically divided into two categories, mountains and plains. The advantage of division is that we can ensure the points in the same cluster having high similarity; the similarity is small in different divisions. The cluster analyzing results reveal the characteristic value for the plain areas (δ1) and the mountain areas (δ2), and based on them the adjacent elevation values between grid points can measure the micro-topographic features. Plain landform, because of having less detail, is gently undulating; the detail of alpine terrain is more complex, the landscape may have dramatically changed in a small region. So, the geomorphology distance autocorrelation is possible clear only in a small area. In the classification thresholds δ1 and δ2, we can judge the current terrain features, and then combined with spatial autocorrelation theory, and compute the elevation values of the center point by the gradient-based distance-weighted algorithm.

4. The Gradient-based Distance-weighted Algorithm

On the basis of K-means clustering analysis for the landforms, we can get the mean value of each cluster which was the threshold to judge terrain features and determined the number of search direction that is involved in computing the discrete points. Getting 4 or 12 discrete points around the new grid calculated the inclination as the weights; the tilt angle is large and the weight on the large, otherwise small. For a flat area (*i.e.* the clustering center value is δ1), because of the less detail of the terrain and the better spatial autocorrelation, we can use fewer points (4 directions) to calculate the elevation value of the center point; for sharp undulating topography (i.e. the clustering center value is δ2), the greater the inclination angle of the area, the bigger of the weights, we use dense points (12 directions) for computing the elevation value to avoid being affected by the local elevation error.

Figure 1 shows the both cases for selecting sampling points in a regular grid. p_1,p_2,p_3,p_4, four adjacent points on the grid, h_1~h_4 is the corresponding elevation values, m_1 is the center point of the grid, and the following m_2 ,m_3 ⋯ is also. The goal is to determine the elevation value of m_i (i=1，2，3⋯n).

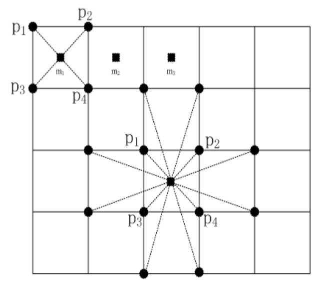


Figure 1. Grid of a Terrain

The relationship between m_1 and its adjacent points p_i can be described by the inclination. The tilt is defined as $\angle h_j m_1 p_j$ (j=1, …, k, k=4 or 12), the tangent value, calculating formula: $\tan\Theta_i=|h_i p_i|/|m_1 p_i|$. The weight of p_i is:

$$w_i = \frac{\tan \theta_i}{\sum_{i=1}^{k} \tan \theta_i} \qquad (1)$$

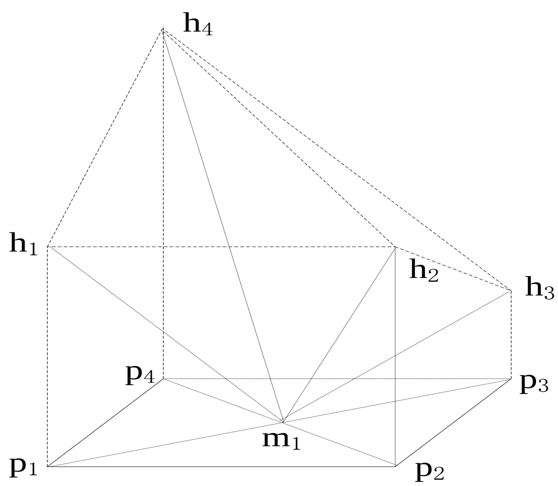


Figure 2. Computing the Center Elevation

As shown in Figure 2, the number of discrete points is decided by the operation of $|p_i-\delta_1| \leq |p_i-\delta_2|$, p_i: nearest points p_1,p_2,p_3,p_4. If the numbers of points meeting the conditions are more than half, the local autocorrelation of terrain is better and we adopt four points to calculate.

On the contrary, the terrain is more intense; we extend the four edges outward, and the distance from the points in the extended edge to the center point is closer than others in grid. So, the extended eight points and the adjacent four points have a strong spatial-relationship.

The functions for computing elevation values as follows:

$$f = \sum_{i=1}^{k} w_i h_i \qquad (2)$$

In formula (2) , f is the elevation of middle points; h_i is the elevation of corresponding point p_i, k=4 or 12.

5. The Processes of Terrain Simplification

According to the results of the processing, we design the algorithm flow of terrain simplification based on spatial autocorrelation is shown in Figure 3.

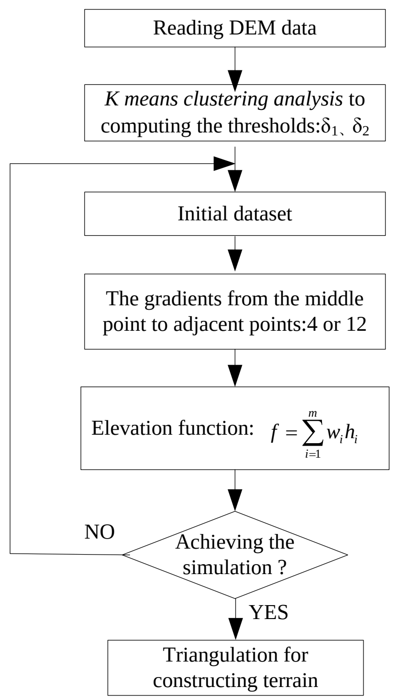


Figure 3. Terrain Simplification Algorithm

The terrain grid is divided by clustering analysis of DEM data, and gets the classification value: δ1 and δ2; secondly, compute the elevation of new points by the gradient-based distance-weighted algorithm; lastly, use triangulation method to construct three- dimensional terrain. The terrain created by new data grid is displayed through OSG platform.

Because the algorithm is carried out in advance of the elevation data clustering analysis, data aggregation can be easily drawn features; these data reflect the characteristics of the landscape that is characteristic terrain rallying point. Slope with an inclination to quantify, but based on local autocorrelation features and terrain data classification results can be adaptively adjusted. Each grid contains new features as data points, so the new grid is a smooth transition from the original grid, both to meet the impact caused by topography, terrain and avoid wrinkles, smoother and more realistic visually.

6. Experiments and Analysis

6.1. Experiment Data

Select the SRTM DEM data (http://www.gscloud.cn/). The data is constituted by a set of uniform regular grid, cell size: 0.00083333°×0.00083333°. The geographic scope of longitude is between 110°~115°, and latitude is between 30°~35°.The column of SRTM data 59, and the row number is 06, the middle-west of Henan province. This region has obvious geographical characteristics, such as plainly areas and less mountainous.

6.2. Error Analysis and Statics

The elevation of fitting points calculated by the above method may be deviation from the real value. Figure 4(a), (b) revealed the elevation deviation between two. The curve of ABCDE in Figure 4(a) represents a real surface of the terrain and B'C'D' is the possible fitting point. Because of the rolling landform, we considered the distance from O' to O'' as the real one, obviously there are some errors. The simplification method should minimize the errors, such as ΔM_b ,ΔM_c and ΔM_d . Then the simulating points can truly reflect the features of the terrain. The formula $H=f(x,y,z,\tan\Theta)+\Delta M$ represents the quantitative relationship between two elevation values, f is the approximation function adopted above, ΔM is the error.

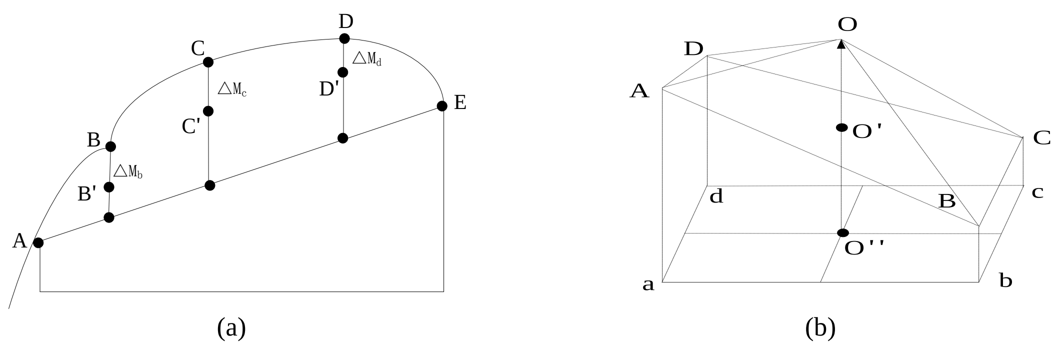


Figure 4. (a) Side View, (b) Three-Dimensional Drawing

The checkpoint method is used to evaluate the accuracy of elevation values. It evaluates the accuracy by computing the value of standard deviation between the points and the fitting points. Considering the characteristic of the data using here, we arranged the checking points evenly distributed on the grid. Then, we computed the relative error of every point in the grid to estimate the effectiveness of the algorithm.

Figure 5 is the distribution of relative error counted. The horizontal axis represents the range of the relative error, and the vertical axis represents the percentage of points within the range. The statistical results show that the errors in new points will be no more than 5% in most case. So, the gradient-based distance-weighted method can be well simulating the elevation point of the terrain surface.

We also use distance weighting method to reconstruct the mesh grid and performed analysis of the errors. The results are shown in table 1: the errors produced by gradient-based distance-weighted method (called solution I) are expressed as $\Delta M1$; the errors produced by distance weighting method (called solution II) are expressed as $\Delta M2$; the expression, such as count ($\Delta M1<\Delta M2$), computed the numbers of error $\Delta M1$ less than $\Delta M2$. Statistical calculation results show that the algorithm in this paper under most circumstances is better than solution II.

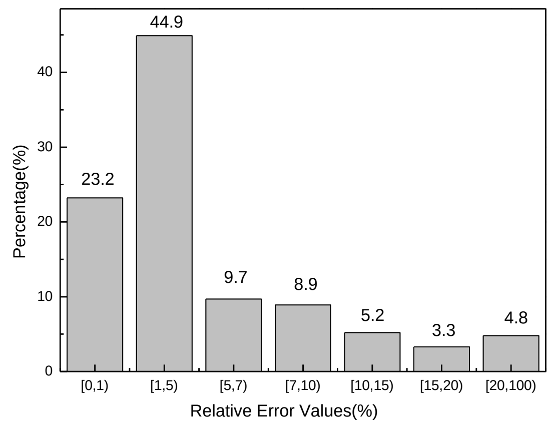


Figure.5 Relative error distribution (6000×6000 grid mesh)

Table 1. Comparison of statistical errors of two solutions

Solutions	Percentage	Description
count(△M1<△M2)	61.046%	The number of points that solution I is better than solution II
count(△M1=△M2)	3.549%	The both are equal
count(△M1＞△M2)	35.405%	The number of points that solution I is worse than solution II

6.3. Curvature Computation

Profile curvature is estimation along the direction of maximum slope; plan curvature is estimation that is perpendicular to the direction of maximum slope. The former describes the rate of change in x direction; the latter describes the rate of change in y direction. If the both values are relatively close in the new grid and old one, the new fitting point is credible. The formulas are as follows:

$$K_r = -\frac{p^2 + 2pqs + q^2 t}{\left(p^2 + q^2\right) + \sqrt{1 + p^2 + q^2}} \quad (3)$$

$$K_n = -\frac{q^2 r - 2pqs + p^2 t}{\left(p^2 + q^2\right) + \sqrt{1 + p^2 + q^2}} \quad (4)$$

In formula (3) and (4) ,p=ӘH/Әx, the rate of change along the x-direction; q= ӘH/Әy, the rate of change along the y-direction; s= $Ә^2$H/ӘxӘy, t= $Ә^2$H/$Әy^2$. Figure 6 is a graph of Kriging interpolation to total curvature raster on original mesh data and reconstructed mesh data, which are handled by the tools of “Curvature and Kriging command” in ArcGIS desktop10.1. It can be seen that the distribution of curvature values is basically consistent.

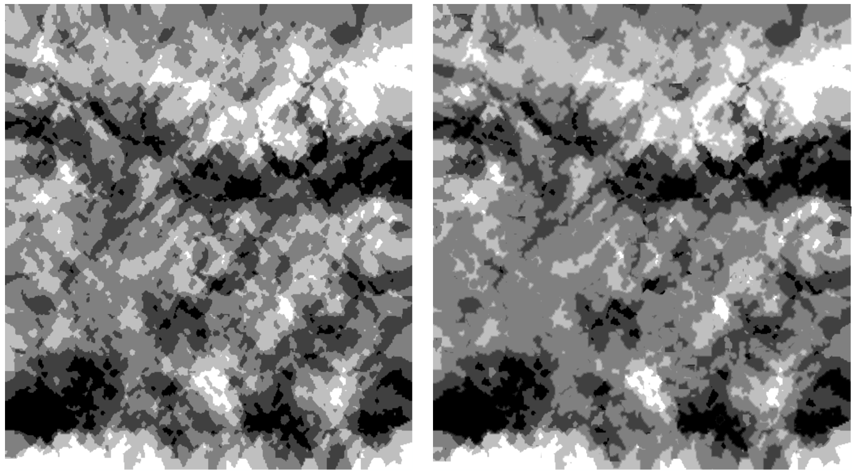
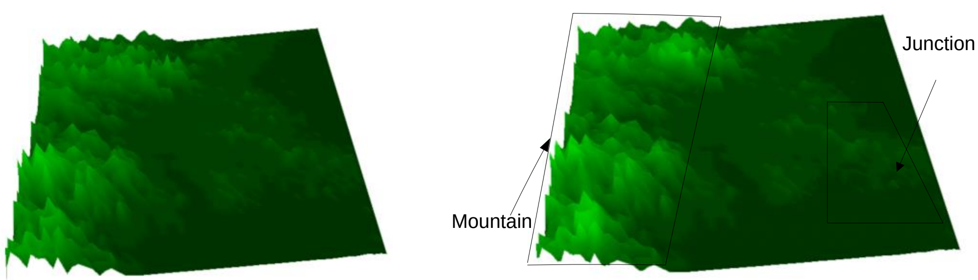


Figure 6. Comparison of total curvature raster of before and after simplification

6.4. Terrain Rendering

Figure 7 is the effect of terrain loaded by OSGsystem. Topographic features keeps consistent with the original image; the reconstructed grid can accurately express the characteristics of the surface, and don't have the cracks, dislocation and other issues. From the figure, we can very well distinguish the mountains and plains. This is due to the terrain topography expression is the most important factor, and the method to calculate the new data points in this article are around 4 or 12 points to the inclination of the center as a weighting factor, considering the high degree of terrain elevation changes impact values, and therefore can be better maintained terrain features. From the above analysis, the relatively large size of the terrain, since the local autocorrelation features elevation values by a new grid to replace the original grid, you can reduce the size of data to some extent, but they can better maintain the terrain features.



(a) (b)

Figure 7. (a) effect of modeling before simplification,(b) effect of modeling after simplification

7. Conclusions and Future Work

In the process of constructing large-scale terrain, the ability to faithfully reshape the landscape's characteristic will have a direct impact on the quality of modeling result, and the size of the data involved in building the terrain will affect the performance of graphic's displaying. In the study of monitoring and warning subjects for the featured fruits, because of the regional distribution regularities of the economic crop, we have designed a three-dimensional visualization module for the purpose of giving users an intuitive impression of the monitoring results. To solve the problem of low displaying performance in 3D visualization modules, we simplify the modeling data based on this algorithm. Not only reduce the amount of data modeling, but also improve the efficiency.

Our future work could be extended to refine the characteristics of the local area, and to increase the number of iterations. So we can improve the performance of the output of terrain model, and the expected size of the data space can be further optimized to improve the calculation of the velocity model, terrain model to optimize the output.

Acknowledgements

The research work was supported by the public service sectors (WMO) research project (GYHY201106024). It was also supported by Industrial Strategic Technology Development Program (10041740) funded by the Ministry of Trade, Industry and Energy (MOTIE) Korea.

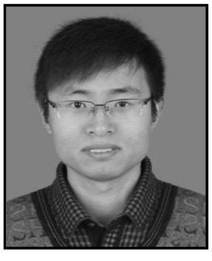
References

[1] H. G. Tang, X. Liu and G. Lv, "Digital elevation model tutorial (Second Edition)", Science Press, Beijing **(2010).**
[2] M. Amor and M. Boo, "Journal of Systems Architecture", vol. 5, no. l**, (2007).**
[3] W. Zheng, "A new mesh simplification algorithm of 3D terrain", 2010 International Conference on Broadcast Technology and Multimedia Communication of IEEE **(2010)**, pp. 283-286.
[4] S. B. Liao and Z. H. Li, "Meteorological Science and Technology", vol. 5, no. 252, **(2004).**
[5] H. Hua and H. Zhao, "Journal of Computer-Aided Design and Computer Graphics", vol. 23, no. 594, **(2011).**
[6] N. Coll and M. Guerrieri, "Journal of computational and Applied Mathematics", vol. 236, no. 1410, **(2011).**
[7] Y. Kho and M. Garland, "User-guided simplification [C]", Proceedings of Symposium on Interactive 3D Graphics, New York: ACM Press, **(2003),** pp. 123-126.
[8] H. D. Zeng, "Geo-information Science", vol. 7, no. 82, **(2005).**
[9] X. Liu, F. Huang and P. Wang, "The principle and method of GIS spatial analysis", Science Press, Beijing **(2008).**

Authors

Yunhua Gu

She received the B.E. degree in the Computer Science and Technology from Shanghai Jiaotong University in 1987, and M.E. degree in the Computer Science and Technology from Southeast University in 2000. Now, she is a professor in the Computer and Software Institute, Nanjing University of Information Science and Technology. Her main research interests are computer network, database, GIS systems. She is a member of ACM and a senior member of CCF.



Xiaoyan Wu

He received the B.E. degree in the Computer Science and Technology from Henan Agriculture University China, in 2010. Now, he is a master degree candidate of Nanjing University of Information Science and Technology, major in computer science and technology. His main research areas are database, GIS systems.

Jin Wang

He received the B.S. and M.S. degree in the Electronical Engineering from Nanjing University of Posts and Telecommunications, China in 2002 and 2005, respectively. He received Ph.D. degree in the Ubiquitous Computing laboratory from the Computer Engineering Department of Kyung Hee University Korea in 2010. Now, he is a professor in the Computer and Software Institute, Nanjing University of Information Science and technology. His research interests mainly include routing method and algorithm design, performance evaluation and optimization for wireless ad hoc and sensor networks. He is a member of the IEEE and ACM.