

A method for hierarchical weighted fitting of regular grid DSM with discrete points

Haoran Guo¹, Weijun Li², Jieke Dong¹, Yansong Duan^{1,3}

¹School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, haoran_guo@whu.edu.cn,
jacky_dot@whu.edu.cn, ysduan@whu.edu.cn

²China Energy Engineering Group Guangdong Electric Power Design Institute Co., Ltd., Guangdong 510700, liweijun@gedi.com.cn

³Hubei LuoJia Laboratory, Wuhan 430079

Keywords: Discrete Points, Regular Grid DSM, Hierarchical Weighted Fitting.

Abstract

A Digital Surface Model (DSM) is a crucial spatial geographic information data used to describe the shape of the earth's surface in Geographic Information Systems (GIS). DSM is the core data used in terrain analysis in GIS. A regular grid DSM is generally generated by interpolating a large number of discrete point clouds. This paper proposes a method of using a hierarchical weighted strategy to fit a regular grid DSM with discrete points. This method uses a pyramid hierarchical strategy to refine the target regular grid from one grid with finer parameters of 3*3, until the nth level (the interval of the grid is equal to the expected interval), and then gradually places the discrete point cloud into the corresponding grid by weighted averaging, and uses the result of this level as the initial value of the next level. This algorithm can avoid the problem of low efficiency in retrieving a large number of discrete point clouds, and the indirect interpolation method not considering the contribution of distant neighboring point clouds. The operation of point cloud data is a stream operation, which does not require consideration of the topological information of point clouds, and has simple operation and no additional memory consumption. It is especially suitable for the production of regular grid DSM with massive point clouds. To verify the effectiveness of this method, the article selected six typical terrain data such as high mountains, mountains, hills, plains, urban areas, and lakes for experiments. The results show that compared with the construct-TIN method for producing DSM, this method has very good processing accuracy and processing efficiency.

1. Introduction

Digital Surface Model (DSM) is an important spatial geographic information data in Geographic Information Systems (GIS) and is the core data for terrain analysis in GIS. It has a wide range of applications in surveying and mapping, geomorphology and geology, engineering construction, military, and other fields (Tang, 2014). DSM mainly includes two forms: regular rectangular grid and Triangulated Irregular Network (TIN) (PETZOLD B, 1999). Among them, the regular rectangular grid DSM is more widely used in photogrammetry due to its simple data structure, easy algorithm implementation, and advantages in spatial operations and storage (Razak, 2011).

In photogrammetry, regular grid DEMs are mainly generated by

interpolating massive and dense matching point clouds or laser point clouds. In actual processing, since large-scale terrain is complex and has large local differences, it is impossible to use a polynomial to fit the entire terrain. Because the fitting accuracy of low-degree polynomials is inevitably poor, and high-degree polynomials may produce unstable solutions. Therefore, global function interpolation is generally not used in DSM interpolation, but local function interpolation is used (Pan, 2023). Local function interpolation divides the entire area into several blocks, and fits each block with different functions, while considering the continuity between adjacent block functions. For the surface that is not smooth or even has a fracture line, further subdivision is required, and smooth or even continuous conditions cannot be used. In addition, point-by-point interpolation methods are widely applied, such as inverse

distance weighting (IDW) method (D Shepard, 1968), moving surface fitting method (SChut, 1976), finite element interpolation method (Ebner, 1978), etc. These algorithms define a local function to fit a certain number of discrete point clouds around the target regular grid point. These methods are flexible and effective, have high accuracy under general conditions, and have simple calculation methods and do not require a large amount of computer memory (Pan, 2023). However, the disadvantages are also obvious. Since it is necessary to search for discrete points around each grid point, the calculation efficiency is extremely low. Especially in large-scale DSM update tasks, it is difficult to meet the requirements. Therefore, some algorithms have been developed to establish block indexes for discrete point clouds, which can improve search efficiency to a certain extent (Chen et al., 2008; Zhang, 2011; Yang, 2013; Chen et al., 2016). However, the calculation process is complex, the complex linked storage structure increases the difficulty of algorithm maintenance, and it is still limited by the size of the survey area and the number of point clouds.

In addition, there is another classic indirect algorithm. This type of algorithm first uses point clouds to construct Triangular Irregular Networks (TIN), and then interpolates the elevation of a regular grid inside the triangle (Dai, 2018). Although indirect algorithms reduce the number of point cloud retrievals and improve efficiency compared to direct interpolation algorithms, the generation of TIN and the determination of the triangle in which the point is located still require a long time. In particular, constructing TIN requires a large amount of computing memory, so indirect algorithms are limited by the number of point clouds. Moreover, since the grid points are only interpolated from the three vertices of the triangle in which they are located, the interpolation accuracy of indirect algorithms is more likely to be affected by a triangle vertex with a large elevation error.

In summary, most of the current image matching direct interpolation algorithms start from ordered target regular grid points and search in massive unordered discrete point clouds. Therefore, the efficiency of searching for discrete point clouds around regular grid points is greatly limited. Meanwhile, for some indirect algorithms, the elevation of regular grid points is only interpolated from the elevations of the three vertices of the triangle in which they are located, so their accuracy is limited. To address these issues, we propose a method for fitting regular

grid DSM with discrete points by a hierarchical weighted strategy. This method fully considers the characteristics of unordered discrete point clouds, only traverses the point clouds in order, places the point clouds into the corresponding grid according to the geographic coordinates, integrates the contributions of neighboring point clouds to the elevation of the grid point through a pyramid structure, and constrains the influence of distant discrete point clouds on the accuracy of the regular grid point through a weighted strategy. Finally, this method achieves fast production of regular grid DSM from discrete point clouds.

2. Methodology

The hierarchical weighted fitting method shares fundamental principles with the inverse distance weighted fitting method, both being rapid distance-weighted fitting algorithms. The basic idea is shown in Figure 1.

The core of this algorithm is to hierarchically place point clouds into a pyramid-structured grid for create DSM from point cloud data. When the point cloud is placed in the first-level grid, the resulting elevation is the average elevation of all point clouds because the grid has only one cell. The second-level grid is based on the grid obtained in the first-level. The grid is first refined (the recommended refinement parameter is 3×3), and then the point cloud is re-placed into the second-level grid, and the initial value in the grid is weighted average (the weight of the initial value in the grid is recommended to be 0.5) to obtain the elevation value of the second-level grid. Then continue to refine the second-level grid. Iterate this process until the interval of the n th-level grid is equal to the expected interval. The specific processing steps are as follows:

Step 1: Establish a pyramid-structure grid. The first-level grid only contains one cell. The first-level grid is divided according to the refinement parameter of 3×3 in the X and Y directions to obtain the second-level grid. Continue to refine until the interval of the N th-level grid is equal to the expected interval (i.e., the resolution of the DSM). The central positions of cells in the N th-level grid corresponds to the regular grid points of the target DSM.

Step 2: Starting from the first-level grid, place discrete points into the grid according to equal weight. Since there is only one

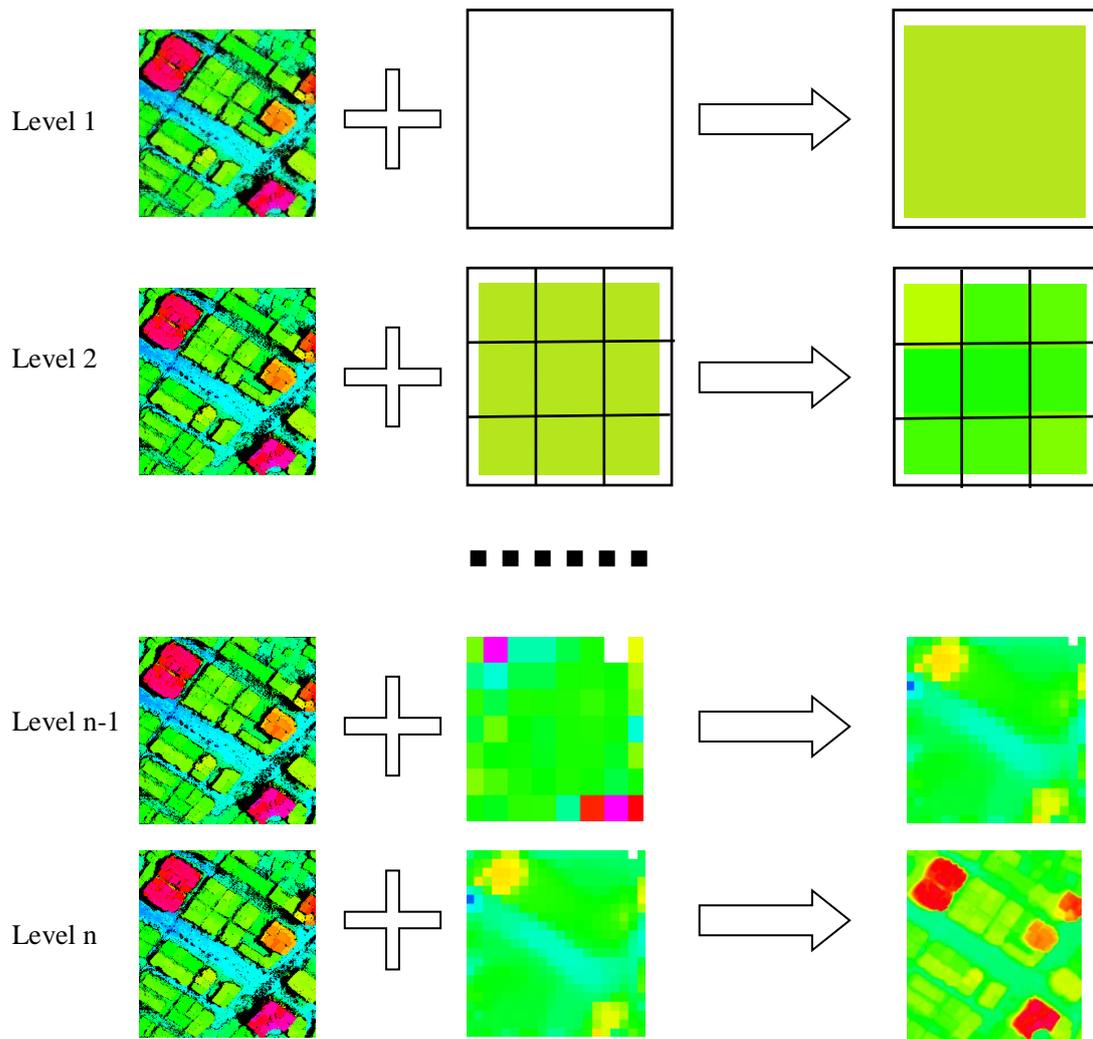


Figure 1. The basic idea of the hierarchical weighted fitting method

cell in the first-level, the elevation of this grid is the average elevation of all discrete points.

Step 3: Divide the first-level grid according to the refinement parameter $P \times P$ to obtain the second-level grid. The second-level grid takes the elevation of the first-level grid as the initial value of the grid elevation. Then, the discrete point cloud is re-placed into the second-level grid and weighted averaged with the initial elevation in the grid (the weight of the initial elevation in the grid is recommended to be 0.5), to obtain the new second-level grid.

Step 4: Continue to refine the grid and re-place all discrete points in the grid to form a new grid. Iterate this process until the N th-level grid is calculated.

Step 5: At this point, the elevation value of each cell in the N th-

level grid is the corresponding grid point elevation of the target DSM.

In the algorithm, the final interpolation value of the elevation of a target grid point is:

$$Z_{grid} = \sum \tilde{Z}_j \quad (1)$$

where $\tilde{Z}_j = \sum_{i=1}^K \frac{Z_j}{S_i * 2^{N+1-i}}$

Z_{grid} = the target grid point's elevation;

\tilde{Z}_j = the contribution value of the j -th discrete point to the grid point;

Z_j = the elevation value of the j -th discrete point;

K represents that from the first-level until the K -level, the j -th discrete point remains in the grid;

S_i = the total number of discrete points in the grid

where the grid point is located at the i -th level grid.

So it can be seen that all point clouds contribute to the elevation of the grid point. The closer a discrete point is to the grid point, the higher its contribution value, and vice versa, the lower its contribution value, even approaching zero.

3. Experiment

3.1 Experimental Data

To verify the effectiveness of the proposed method, the paper selected six typical terrain point clouds, including high mountains, mountains, hills, plains, urban areas, and lakes, for experiments. The information of each data is shown in Table 1. The perspective display of six typical terrain point clouds is shown in Figure 2.

Data	Number of points	Area (Km ²)	Density (points/m ²)	Maximum elevation (m)	Minimum elevation (m)
(a)high mountains	682313	0.707	16.41	1061.47	441.019
(b)mountains	15177123	7.073	3.75	712.851	122.851
(c)hills	12588322	1.479	22.58	341.945	111.289
(d)plains	5670297	0.793	9.21	124.594	107.765
(e)urban areas	2719005	0.712	3.97	207.167	48.2323
(f)lakes	1258340	0.252	6.10	143.25	122.484

Table 1. Experimental data information

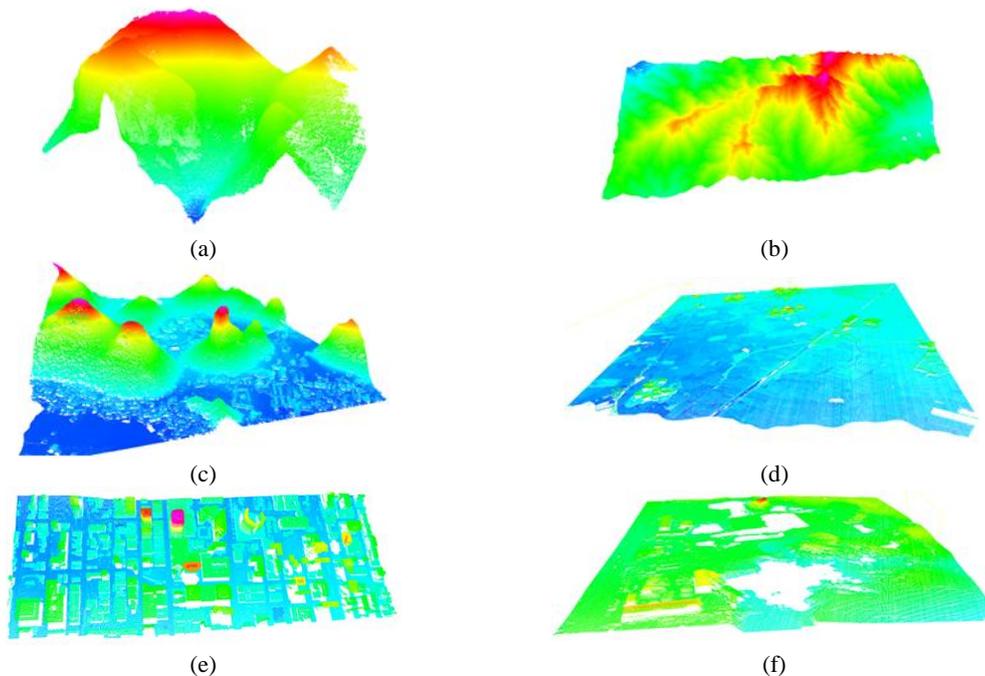


Figure 2. Point cloud data for 6 typical terrains: (a) high mountains, (b) mountains, (c) hills, (d) plains, (e) urban areas, (f) lakes

3.2 Experimental Scheme

To verify the efficiency and fidelity of the method proposed in this paper, a comparative experiment was designed. Two methods were used: the hierarchical weighted method in this paper (hereinafter referred to as Grd) and the indirect method of constructing TIN to generate DSM (hereinafter referred to as Tin). Six typical terrain point cloud data were used to produce DSMs with the same resolution. Then, the two methods were compared based on two criteria: processing speed and fidelity.

Figure 3 shows the experimental scheme.

The evaluation criteria for the two methods are as follows:

Criterion 1: Processing speed

Method: Generate DSMs with the same resolution using the two methods for each set of point cloud data, and record the processing time for each method.

Criterion 2: Fidelity

Method: Use the XY coordinates of each point in the point cloud to interpolate Z' in the resulting DSM. Calculate the difference dZ between Z' and the point cloud coordinate Z . Obviously, the smaller the difference, the higher the fidelity of the method used.

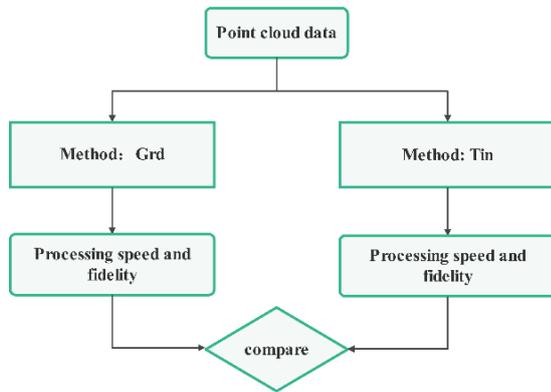


Figure 3. Experimental scheme

3.3 Experimental Result

The resulting DSMs of six typical terrain data generated by two methods are shown in Figure 4 - Figure 9.

As shown in Figure 4 and Figure 5, the Grd method and the Tin method have comparable effects in displaying the details of mountainous terrain and serpentine roads. In Figure 6, the DSMs generated by the two methods show comparable effects for houses. For the top of the tower, the Tin method performs better than the Grd method. This is because Tin contains more terrain feature points, which is more conducive to recovering the top of the tower. As shown in Figure 7, for flat terrain, both methods can preserve the geomorphic features well. In Figure 8, for the urban data set, since there are no buildings with pointed roofs, the Grd method and the Tin method produce similar effects. Comparing Figure 9(b) and Figure 9(d), it can be seen that the Grd method is superior to the Tin method. In the DSM generated by the Grd method, the surface of the lake in the upper left corner is smoother. At the bottom of Figure 9(d), there are many obvious triangles, and the Tin method does not show the original topography.

Overall, the DSMs generated by two methods have comparable effects. The Tin method is more conducive to preserving sharply changing geomorphic features. The Grd method is more suitable for the production of DSMs in flat areas.

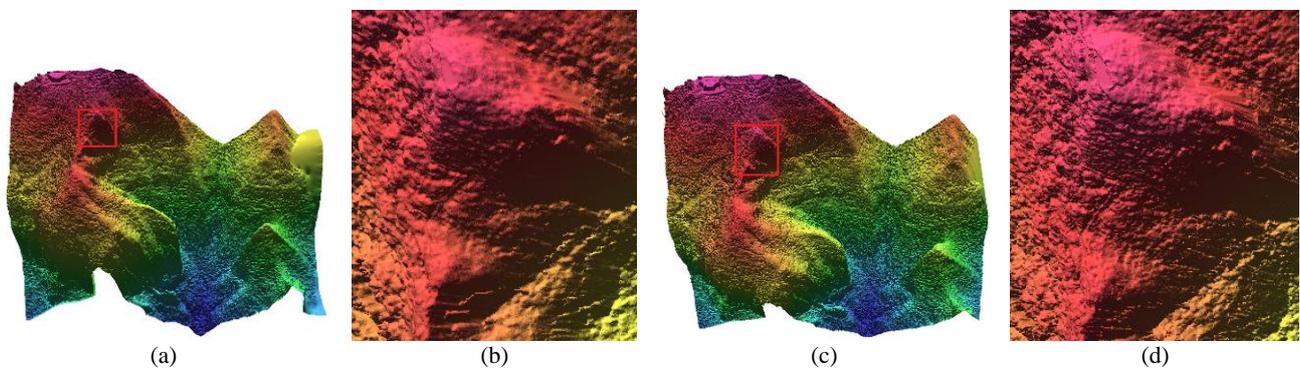


Figure 4. Resulting DSMs of high mountains: (a) DSM (Grd), (b)Partial image of DSM(Grd), (c) DSM (Tin), (d) Partial image of DSM(Tin)

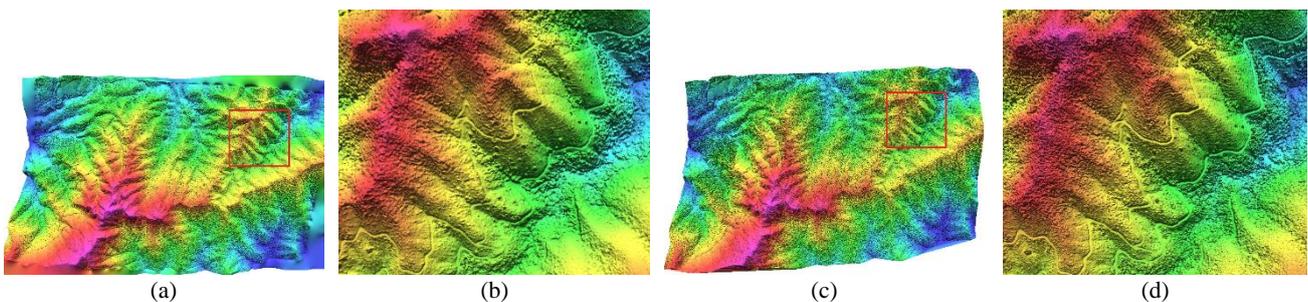


Figure 5. Resulting DSMs of mountains: (a) DSM (Grd), (b)Partial image of DSM(Grd), (c) DSM (Tin), (d) Partial image of DSM(Tin)

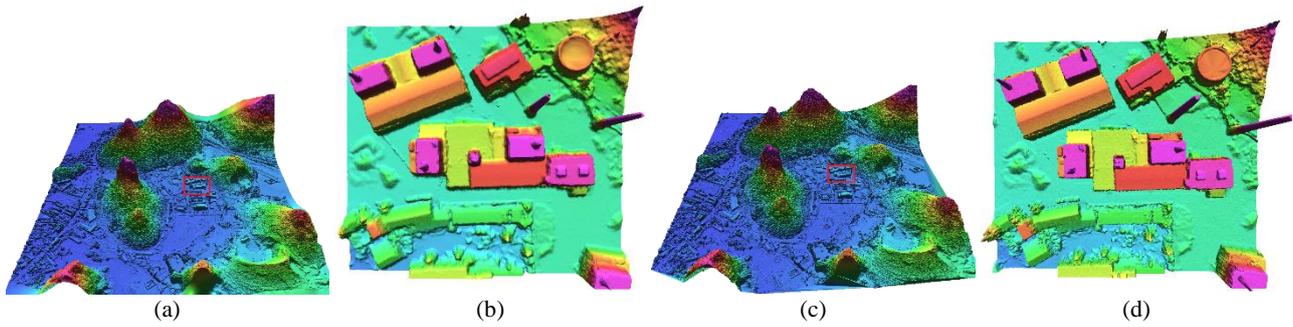


Figure 6. Resulting DSMs of hills: (a) DSM (Grd), (b)Partial image of DSM(Grd),
 (c) DSM (Tin), (d) Partial image of DSM(Tin)

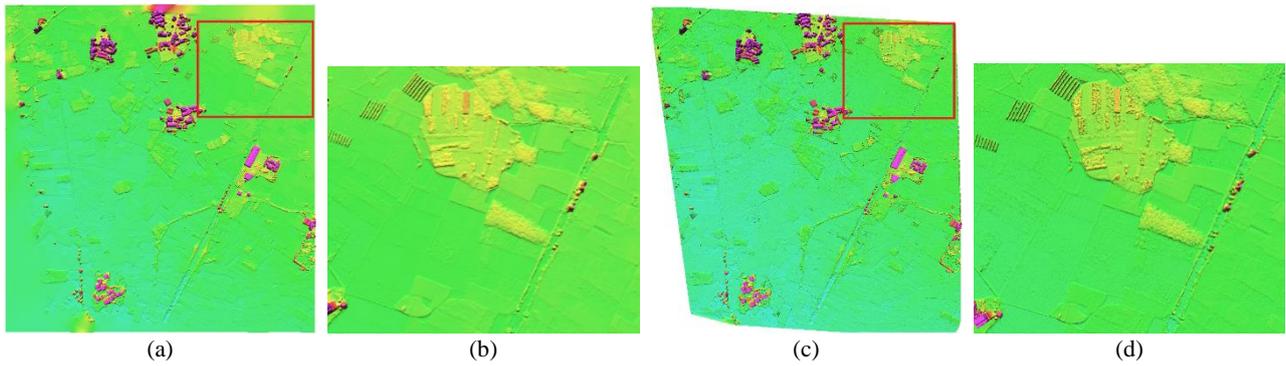


Figure 7. Resulting DSMs of plains: (a) DSM (Grd), (b)Partial image of DSM(Grd),
 (c) DSM (Tin), (d) Partial image of DSM(Tin)

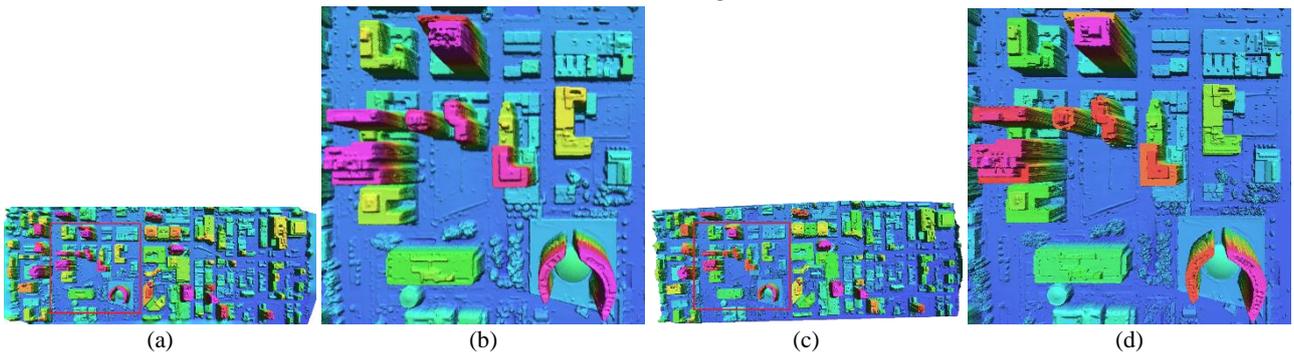


Figure 8. Resulting DSMs of urban areas: (a) DSM (Grd), (b)Partial image of DSM(Grd),
 (c) DSM (Tin), (d) Partial image of DSM(Tin)

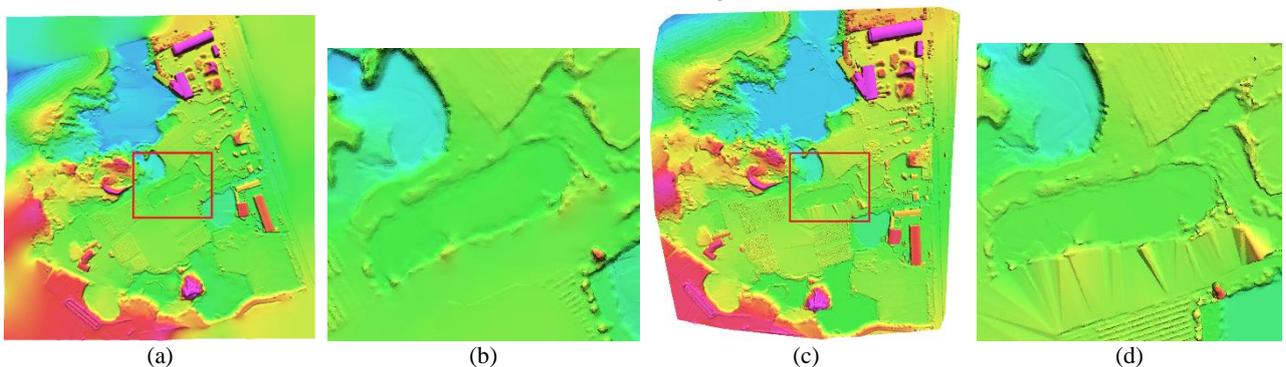


Figure 9. Resulting DSMs of lakes: (a) DSM (Grd), (b)Partial image of DSM(Grd),
 (c) DSM (Tin), (d) Partial image of DSM(Tin)

The first evaluation indicator is processing speed. The processing times of the two methods are shown in Table 2. By

comparison, it can be seen that the Grd method in this paper takes about half the time of the Tin method. This is mainly

because constructing TIN and determining the triangle where the grid point is located take more time. Using the method

proposed in this paper can improve the efficiency of generating DSMs from discrete point clouds.

Data	The Number of points	DEM resolution(m)	Processing time of Grd (s)	Processing time of TIN (s)
high mountains	682313	0.25	35	110
mountains	15177123	0.45	55	116
hills	12588322	0.20	50	110
plains	5670297	0.30	9	18
urban areas	2719005	0.45	5	9
lakes	1258340	0.35	4	5

Table 2. Process time of two methods

The second evaluation indicator is fidelity. Calculate dZs for all points in the six sets of point cloud data. And then the distribution of dZ values is statistically analyzed. Figure 10 shows the violin plot of the dZ distribution for six sets of point cloud data. The body of the violin plot shows the frequency distribution, and the wider it is, the higher the frequency near a certain value. From Figure 10, it can be seen that the dZ distributions obtained by the Grd method and Tin method are approximately the same. The fidelity of the six sets of point cloud data can be ranked from high to low as follows: lakes > plains > hills > mountains > high mountains > urban areas. Table 3 lists the mean, variance, and percentages of dZ absolute values within 1m, between 1m and 5m, and above 5m for each set of point cloud data. According to Table 3, the means of dZ obtained using both methods are close to 0, the variances are similar, and most of the dZs are within 1m. Based on Figure 10 and Table 3, it can be concluded that the fidelity of the Grd method and Tin method is comparable.

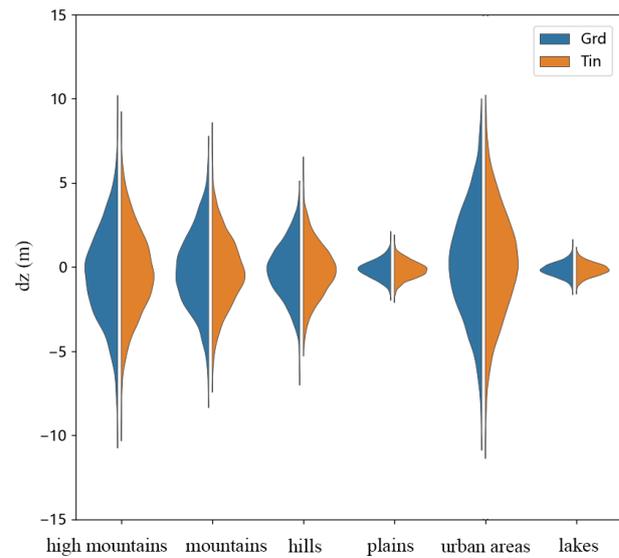


Figure 10. The violin plot of the dZ distribution

Terrain	Method	Mean (m)	Std (m)	$ dZ < 1m$	$1m < dZ < 5m$	$ dZ > 5m$
high mountains	Grd	-0.2418	2.5293	30.48%	64.56%	4.96%
	Tin	-0.0882	2.3808	32.08%	64.44%	3.48%
mountains	Grd	-0.0184	2.1104	35.76%	62.18%	2.06%
	Tin	-0.1406	1.9370	40.32%	58.48%	1.20%
hills	Grd	-0.0359	1.5692	48.20%	51.66%	0.14%
	Tin	-0.0302	1.5087	50.80%	49.16%	0.04%
plains	Grd	0.0216	0.4903	95.62%	4.38%	0.00%
	Tin	0.0162	0.4914	96.08%	3.92%	0.00%
urban areas	Grd	0.4006	2.9674	25.96%	64.10%	9.94%
	Tin	0.2986	2.9068	26.58%	64.44%	8.98%
lakes	Grd	0.0145	0.3843	99.01%	0.99%	0.00%
	Tin	0.0151	0.3606	99.36%	0.64%	0.00%

Table 3. Statistical table for the distribution of dZs

4. Conclusion

This paper proposes a method for fitting regular grid DSM with discrete points by a hierarchical weighted strategy to address the low efficiency of existing technologies in retrieving massive

discrete point clouds and the issue of indirect interpolation methods not considering the contributions of distant neighboring point clouds. The paper conducts experiments using six types of typical terrain data, including high mountains, mountains, hills, plains, urban areas, and lakes. The results

demonstrate that the proposed method in this paper to fit regular grid DSM improves processing efficiency by more than twice compared to the indirect method of constructing TIN to generate DSM through interpolation, while maintaining comparable data fidelity. Therefore, we believe that the hierarchical weighted method in this paper has better prospects for applications. Additionally, if computing resources are limited, such as mobile devices, edge computing devices, etc., this method has better performance.

Acknowledgements

This study was supported by National Key Research and Development Program of China, No. 2023YFB3905704.

References

- Chen, G., Zhang, X., Zhang, M., Shi, H., 2008. Study of the Organization of LiDAR Points Cloud Data Based on Quadtree. *Bulletin of Surveying and Mapping*, (11):21-23.
- Chen, M., Wan, Y., Tian, S., et al., 2016. A Method of Organizing Point Clouds Based on Linear KD Tree. *Bulletin of Surveying and Mapping*, (01):23-27.
- Dai, L., Niu, J., Zhang, L., et al., 2018. Analysis of DEM Production Process Based on Airborne LiDAR. *Geospatial Information*,16(10):44-46+57+10.
- D Shepard, 1968. A two-dimensional interpolation function for irregularly-spaced data. *Acm National Conference*,517-524.
- Ebner, H., Reiss, P., 1978. Height Interpolation by the Method of Finite Elements. Presented Paper, ISP-Commission III-Symposium, Moscow. *Nachrichten aus dem Karten- und Vermessungswesen*, Reihe II, Heft 36,79-94.
- Pan, L., et al., 2023. *Photogrammetry*. Wuhan: Wuhan University Press.
- PETZOLD B, REISS P, STÖSSEL W, 1999. Laser scanning-Surveying and mapping agencies are using a new technique for the derivation of digital terrain models. *ISPRS Journal of Photogrammetry and remote Sensing*, 54(2/3): 95-104.
- Razak, K. A., M. W. Straatsma, C. J. Van Westen, J. P. Malet, S. M. De Jong, 2011. Airborne Laser Scanning of Forested Landslides Characterization: Terrain Model Quality and Visualization. *Geomorphology (Amsterdam, Netherlands)*, 126(1-2): 186-200.
- Schut G. H., 1976. Review of Interpolation Methods for DTMS. Invited paper, ISP. Comm. III, Helsinki.
- Tang, G., 2014. Progress of DEM and digital terrain analysis in China. *Acta Geographica Sinica* ,69(09):1305-1325.
- Yang, J., 2013. 3D R-Tree Point Clouds Data Indexing Method Based on Plane Fitting. *Geomatics and Information Science of Wuhan University*, 38(11):1313-1316.
- Zhang, H., 2011. The Organization and Visualization of Point Cloud Data Based on the Octree. *Journal of Taiyuan Normal University (Natural Science Edition)*, 10(03):128-132.