information

Article

Research on Building DSM Fusion Method Based on Adaptive Spline and Target Characteristic Guidance

Jinming Liu [1], Hao Chen [2,*] and Shuting Yang [2]

[1] Institute of Defense Engineering, AMS, PLA, Beijing 100036, China; orlandorothfeld@gmail.com
[2] School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150006, China; 21B905002@stu.hit.edu.cn
* Correspondence: hit_hao@hit.edu.cn

Abstract: In order to adapt to the actual scene of a stereo satellite observing the same area sequentially and improve the accuracy of the target-oriented 3D reconstruction, this paper proposed a building DSM fusion update method based on adaptive splines and target characteristic guidance. This method analyzed the target characteristics of surface building targets to explore their intrinsic geometric structure information, established a nonlinear fusion method guided by the target characteristics to achieve the effective fusion of multiple DSMs on the basis of maintaining the target structural characteristics, and supported the online updating of DSM to ensure the needs of practical engineering applications. This paper presented a DSM fusion method for surface building targets and finally conducted DSM fusion experiments using typical urban area images of different scenes. The experimental results showed that the proposed method can effectively constrain and improve the DSM of buildings, and the integrity of the overall construction of the target 3D model structure was significantly improved, indicating that this paper provides an effective and efficient DSM constraint method for buildings.

Keywords: remote sensing imagery; 3D reconstruction; DSM; adaptive splines

Citation: Liu, J.; Chen, H.; Yang, S. Research on Building DSM Fusion Method Based on Adaptive Spline and Target Characteristic Guidance. *Information* **2021**, *12*, 467. https://doi.org/10.3390/info12110467

Academic Editor: Stefano Berretti

Received: 9 October 2021
Accepted: 9 November 2021
Published: 10 November 2021
Corrected: 7 January 2022

1. Introduction

Three-dimensional (3D) reconstruction is a research hotspot in the application of computer-aided engineering technology for product and engineering design [1–3]. Three-dimensional (3D) reconstruction is a new technology that integrates knowledge in the fields of machine vision, image processing, and computer cartography and has characteristics that cannot be compared with 2D pictures. Traditional remote sensing 3D reconstruction methods usually aim at large-area and large-scale scene reconstruction and tend to realize the model construction of the overall structure of the target from a global perspective, commonly without focusing on the local structure of the target, and it is difficult to ensure the structural integrity and accuracy of the target with complex structure [4–6].

In the field of target 3D reconstruction, the main data sources at this stage include digital surface model (DSM) [7], digital elevation model (DEM), digital terrain model (DTM), and airborne LiDAR data [8] generated from aerial imagery and satellite high-resolution image interpretation. In the past decades, researchers have proposed many methods to implement target 3D model construction based on DSM or LiDAR data, which can be summarized into three categories: model-driven, data-driven, and hybrid-driven methods [9–11].

Model-driven methods start with the construction of a library of predefined primitives containing the basic model of the target, based on which the target 3D model construction is achieved by matching the structure between the predefined model and the input data [12]. Model-driven methods are commonly applied to the 3D reconstruction process of models whose targets are projected as rectangles in the 2D plane. Vosselman et al. [13] used

geometric invariant moment features for primitive model matching and estimated the parameters of the model primitives to achieve 3D model construction. Subsequently, Brenner et al. [14] proposed a splitting rule for target shape and structure and realized the construction of a three-dimensional model of the target with a more complex structure. The model construction method proposed by Henn et al. [15] used the SVM algorithm to statistically classify the point clouds of buildings within the rectangularly divided area of the target bottom surface and the RANSAC algorithm to fit the model parameters to achieve the estimation of the target model structure parameters to obtain the final target construction results. Lafarge et al. [16] combined DSM data and the extracted 2D contours of the target for 3D model reconstruction, which is based on the concept of model splitting and stitching, using the Gibbs model to control the combination mode between each primitive model of the target, using Markov Chain Monte Carlo (MCMC) to find the optimal combination mode between model primitives, and estimating the structural parameters of each primitive model, so as to realize the 3D reconstruction of the overall structural model of the target. Huang et al. [17] proposed a generalized model construction method that treats the construction of the target as a top-down process, predefines the parametric description and combination method of the primitive models in the model library, and uses the Reversible Jump Markov Chain Monte Carlo (RJMCMC) technique combined with the structural transformation kernel function to achieve 3D model construction of targets with different structures. However, it is difficult to construct targets with complex structures based on model-driven target 3D reconstruction methods. Due to the complex and variable target geometric structure, it is difficult to parametrically describe the 3D model of targets with different degrees of complexity by a unified method, which will affect the completeness and accuracy of the overall structural model construction of the target without the constraints of the global structure.

In recent years, data-driven model construction methods have been widely used in the field of 3D reconstruction due to their higher flexibility [18–20]. The data-driven model construction method first uses segmentation methods to obtain different adjacent subsurfaces of the top surface of the target and improves the regularization of the overall structure of the target by adjusting the structure of adjacent points and lines between the subsurfaces to obtain the final model construction results. Verma et al. [21] proposed a data-driven model construction method, which introduced a seed region selection method based on the similarity of local features based on the traditional segmentation method of region growth to achieve the segmentation of the top surface of the target building by effectively controlling the range of region growth through the constraints of the structural features of local regions. Finally, the results of the 3D reconstruction of the model are obtained by combining the boundary points of the target top surface and the connection topology of the ridge line features. Chauve et al. [22] proposed a region growing method based on topological constraints on the target top surface for segmentation, which added structural constraints to the topological connections of the target top surface to the traditional region growing algorithm, thus enabling more accurate acquisition of information on each subplane included in the top surface of the target, which improved the completeness of the target model construction. Sampath et al. [23] used the fuzzy K-mean clustering technique to segment the target top surface and obtain the boundary and ridge features of each subplane and adjusted the connection between each vertex, as well as the line features according to the predefined normalized structure criterion to achieve the model construction of the overall structure of the target. Kim et al. [24] used the level-sets approach for model construction, which firstly constructed an energy function based on multitemporal level sets and obtained the segmentation results of the target top surface according to the minimization process of solving the energy function and then obtained the contours and ridges of the target top surface according to the output of the zero-level energy function and finally obtained the final model 3D reconstruction results according to the boundaries of the adjacent subsurfaces of the top surface, the ridges, and the topological connection relationships between the vertices. The limitation of the data-driven approach

is that the quality of model construction is too dependent on the segmentation results of the target top surface. When there are certain noise or singularities in the data source, the segmentation for the target top surface usually has the phenomenon of over-segmentation or undersegmentation, resulting in the final construction results losing local structural features, and thus affecting the structural integrity of the target construction model.

In order to further improve the stereo model construction, a hybrid stereo model construction method combining data-driven and model-driven has been widely used [25,26], which mainly combines the advantages of model-driven and data-driven methods and firstly constructs a library of well-formed building primitive models based on the model-driven concept and predefines the parameterization of different types of model primitive description methods. Secondly, the data-driven method is used to detect structural features such as boundaries and ridges of the target, and finally, the 3D reconstruction of the target is realized by the constraints of the connection rules of these basic structural elements. Xiong et al. [27] proposed a dictionary library based on the structured cartographic description of the model, which stores the structured cartographic drawings for each type of structure, and the dictionary library has representative error subdrawings and corresponding structure correction subdrawings, based on which the structural integrity and accuracy of the model 3D reconstruction are further improved by the techniques of cartographic matching and structure correction.

In summary, traditional 3D reconstruction methods are aimed at building large regional terrain or urban scenes, and commonly neglect the fine construction of the local structure of the target. In the process of 3D reconstruction for the target, there is a higher demand for the completeness and accuracy of the overall construction of the target 3D model structure, and more focus on the high-precision construction of the local structure of the target, which are the challenges faced by the traditional 3D reconstruction methods.

The remainder of this paper is organized as follows. Section 2 details the specific workflow of research, including building geometry characteristics, adaptive spline-based DSM fusion guided by building geometric characteristics. Section 3 is a comparative analysis of experimental results. Discussions and conclusions are presented in Sections 4 and 5, respectively.

2. Methodology

In this paper, we first analyzed the target characteristics of the surface building target to explore the geometric structure information inherent in the building, established a nonlinear fusion method guided by the target characteristics to achieve the effective fusion of multiple DSMs on the basis of maintaining the target structural characteristics, proposed the DSM fusion method for the surface building target, and finally conducted DSM fusion experiments using typical urban area images of different scenes.

2.1. Building Geometric Structure Characteristics

A typical building is composed of beam–slab–column elements, and its roof is mostly composed of colorful steel sheets or modified bitumen. In this paper, we considered exactly these types of buildings. From the geometric point of view, the roof shape of a typical building is a polyhedral structure, a representative example of which is shown in Figure 1 (derived from reference [17], titled "A generative statistical approach to automatic 3D building roof reconstruction from laser scanning data").

According to Figure 1 (derived from reference [17], titled "A generative statistical approach to automatic 3D building roof reconstruction from laser scanning data"), it can be found that typical building roof shapes are diverse and contain flat roofs, pitched roofs, double-slope roofs, four-slope roofs, and multiwave folding slab roofs. Although the specific manifestations of roof shapes are diverse, they can be considered to consist of roof panels with different inclination angles. For this reason, it is necessary to construct a unified mathematical description of building roof characteristics to accommodate the different roof geometric structures of typical buildings.

For building roofs with different morphologies, this paper constructed adaptive 2D spline functions to describe building geometry universally. For the 2D spline function, each roof panel of a building can be considered as a linear spline with specific parameters in 3D space, the flat roof of a building can be considered as a 0th-order 2D spline, the pitched roof with different inclinations can be represented as a 1st-order spline with different parameters, and the roof of the whole building can be considered as a complex and diverse roof shape with linear splines connected by ridge lines with different parameters, as shown in Figure 2.

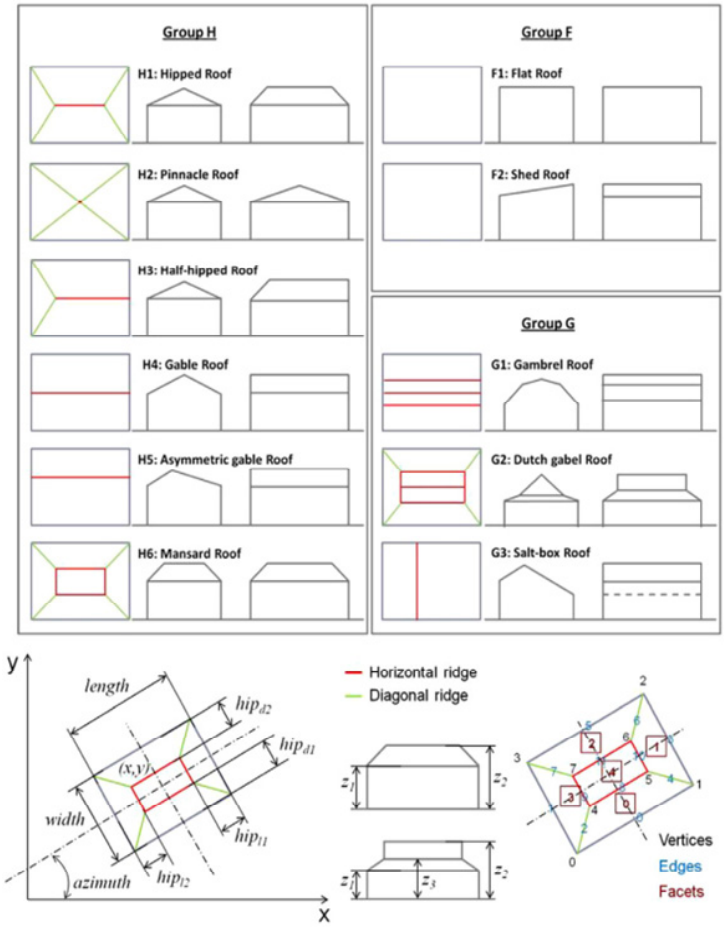


Figure 1. Representative building roof shapes [17].

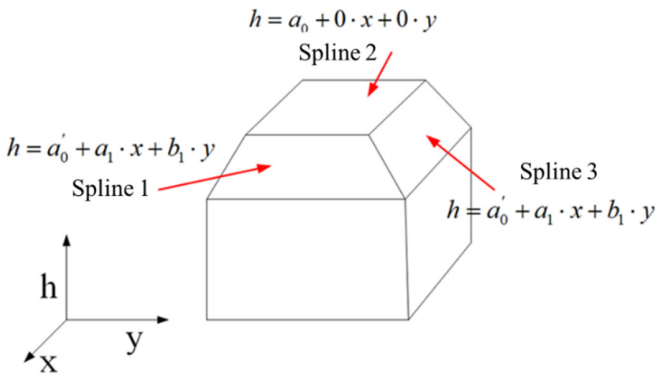


Figure 2. Four-slope-roofed building spline function description.

As can be seen from Figure 2, each roof panel of a four-slope building can be considered as an inclined surface of different angles in 3D space, and the top flat roof panel can be represented as a 0th-order spline, while the four sloping roof panels on the side can be represented as 1st-order splines with different parameters. Connecting the spline functions through the ridge line forms the roof geometry of the building, which means that the spline functions can accurately describe a typical building roof.

For typical flat roofs, pitched roofs and typical buildings with flat or pitched roof panels such as multiwave folding roofs all can be accurately described using the spline function, and the conclusion can be summarized as follows.

Theorem: For a rectangular interval defined in the subspace (x, y) of the right-angle coordinate system (x, y, h), there exists a set of linear spline functions $S_i : I_i \to \mathbb{R}\ 1 \leq i \leq N$ that can accurately fit the roof geometry of typical flat roofs, pitched roofs, and multiwave folding roofs with flat or pitched roof panels under a specific interval division $I_i|_{i=1}^{N}$, and satisfying $I_i \cap I_j = \varnothing, i \neq j, \underset{i}{\cup} I_i = I$ conditions.

Note that the spline function is a segmental function. Only by dividing $I_i|_{i=1}^{N}$ in the appropriate interval can the specific building roof structure be accurately described. The spline function of each region constitutes each roof panel of the building, the intersection line between the intervals corresponds to the ridge line of the building, and for buildings with different roofs, the location and direction of the ridge line are not fixed. In order to characterize the different directions and positions of the roof ridge lines of different structural buildings, the building area needs to be gridded at different scales to approximate the roof ridge lines of different buildings. The gridding of the building area at different scales is shown in Figure 3.

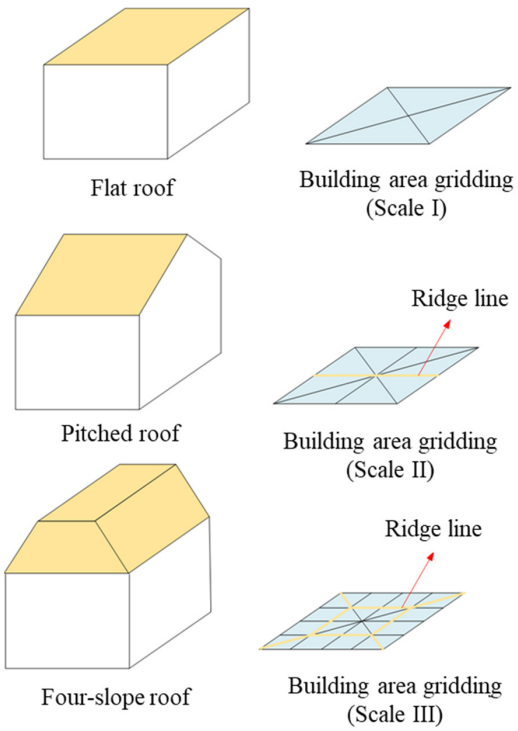


Figure 3. Multiscale gridding of building areas.

It can be seen from Figure 3 that the different scales of meshing of the rectangular area of the building can be adapted to buildings with different roof structures. Specifically, a flat-roofed building can be represented as a 0th-order spline defined by a rectangular grid, a pitched-roofed building can be represented as two 1st-order splines defined by two rectangular grids (separated by the projection of the ridge line at (x, y)), and a four-slope roof can be represented as five 1st-order splines defined by four trapezoids with

one rectangular grid (separated by the projection of the ridge line at (x, y)). Therefore, the spline function only needs a suitable grid division so that the grid line and the ridge line in the (x, y) projection line coincide to ensure that the spline function can accurately describe the corresponding structure of the building. The gridding of scales I, II, and III in the figure can be adapted to flat-roofed buildings, pitched-roofed buildings, and four-slope-roofed buildings. It can be inferred that the denser the gridding, the more accurate the spline function can be in describing the complex roof shape of the building.

2.2. Analysis of Fitting Error of Adaptive Spline to Building Target

In practice, building targets have a wide variety of roof shapes, which means that it is difficult to prespecify the exact grid to achieve an accurate description of arbitrarily shaped roofs. In order to complete the meshing of the building area more effectively, the fitting error of the spline function to the typical building roof shape is analyzed.

For any given building roof structure, that is, any n^{th}-order differentiable function $F : I \to \mathbb{R}$ defined in the rectangular interval I in the subspace (x, y) of the right-angle coordinate system (x, y, h), for any given positive number ε, there exists a sufficiently dense interval division $I_i|_{i=1}^{N}$ that satisfies $I_i \cap I_j = \varnothing, i \neq j, \underset{i}{\cup} I_i = I$ so that the linear spline function $S_i : I_i \to \mathbb{R}\ 1 \leq i \leq N$ interpolates any noncollinear three points of the building roof panel, and the elevation error at any location of the spline function and the building roof structure is less than ε.

Proof: First consider the spline function $S_1 : I_1 \to \mathbb{R}$ in the interval I_1. If it is able to interpolate any noncollinear three points (x_1, y_1, h_1), (x_2, y_2, h_2) and (x_3, y_3, h_3) of the building roof, the parameters of this linear spline $a_0 + a_1 x + b_1 y = h, (x, y) \in I_1$ can be calculated by the following equation.

$$\begin{bmatrix} 1, x_1, y_1 \\ 1, x_2, y_2 \\ 1, x_3, y_3 \end{bmatrix} \times \begin{bmatrix} a_0 \\ a_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} \tag{1}$$

Since a unique plane (i.e., the plane described by the linear splines) can be determined through any noncollinear three points, the location of any point on the plane is an affine combination of the three points, which can be expressed as the following equation.

$$S_1\left(\sum_{i=1}^{3} l_i \cdot [x_i, y_i]\right) = \sum_i l_i \cdot F(x_i, y_i) \sum_{i=1}^{3} l_i \cdot [x_i, y_i] \in I_1 \tag{2}$$

To calculate the elevation difference between the spline function and the building roof function, Equation (3) is expanded using Taylor's formula.

$$\begin{aligned} S_1\left(\sum_{i=1}^{3} l_i \cdot [x_i, y_i]\right) &= \sum_i l_i \cdot \Big(F(x_0, y_0) + \nabla F(x_0, y_0)^T [x_i - x_0; y_i - y_0] + \\ &\quad 0.5 \times [x_i - x_0; y_i - y_0]^T G(F(x_c, y_c))[x_i - x_0; y_i - y_0]\Big) \\ F\left(\sum_{i=1}^{3} l_i \cdot [x_i, y_i]\right) &= F(x_0, y_0) + \nabla F(x_0, y_0)^T \left(\sum_{i=1}^{3} l_i \cdot [x_i - x_0; y_i - y_0]\right) + \\ 0.5 \times &\left(\sum_{i=1}^{3} l_i \cdot [x_i - x_0; y_i - y_0]\right)^T G(F(x_c, y_c)) \left(\sum_{i=1}^{3} l_i \cdot [x_i - x_0; y_i - y_0]\right) \end{aligned} \tag{3}$$

Notice that $[x_i - x_0; y_i - y_0]|_{i=1}^{3} \leq l_{len}$, l_{len} is determined by the size of the spline function grid, so the upper limit of the deviation between the actual elevation of the building roof at any location and the elevation fitted by the spline function is ${l_{len}}^2 \times C$, C is defined in the following equation and can be approximated as a constant in general. Therefore, the elevation deviation is inversely proportional to the square of l_{len}, that is, the denser the mesh, the smaller the maximum elevation deviation. For any given elevation error ε, when the grid is sufficiently dense, that is, $l_{len} \leq \sqrt{\frac{\varepsilon}{C}}$, the maximum elevation error

is less than ε. This means that a sufficiently dense grid division enables the spline function to fit an arbitrary function accurately.

$$C = \frac{1}{l_{len}{}^2} \begin{bmatrix} x_1 - x_0, y_1 - y_0 \\ x_2 - x_0, y_2 - y_0 \\ x_3 - x_0, y_3 - y_0 \end{bmatrix}^T \Big([i_1; i_2; i_3] G(F(x_c, y_c)) [i_1; i_2; i_3]^T + diag(i_1, i_2, i_3) G(F(x_c, y_c)) diag(i_1, i_2, i_3)^T \Big) \begin{bmatrix} x_1 - x_0, y_1 - y_0 \\ x_2 - x_0, y_2 - y_0 \\ x_3 - x_0, y_3 - y_0 \end{bmatrix} \tag{4}$$

2.3. DSM Fusion Guided by Building Geometric Characteristics Based on Adaptive Spline

The spline function with sufficiently dense meshing can accurately describe buildings with arbitrary roof shapes, including not only typical flat, pitched, and multiwave roofs, but also unconventional buildings such as curved and spherical roofs. From an architectural point of view, the design of the roof structure of a building has to consider the waterproof performance and material shaping, which makes the roof of a building mostly flat and pitched. When the roofs are flat and pitched, the spline function with dense meshing can effectively fit the roofs, but the local deformation of the building generated by the inaccurate elevation in DSM will also be fitted by the spline function. In order to maintain the building roof structure as much as possible and avoid the introduction of inaccurate elevation information in the DSM, this paper constructed an adaptive spline-based DSM fusion method guided by the geometric structure characteristics of the building and established an adaptive-meshing spline function to mine the structural information according to the complexity of the building roof structure so as to guide the effective fusion of multiple DSMs and output the building DSM with high accuracy while maintaining the target structural characteristics.

First, the building segmentation method [28] is used to determine the building target in the image, and then the elevation value of the building target in the n groups of DSMs is determined according to the results of the stereo solution. For n DSMs, the points of the building area in the i^{th} DSM are represented as a set $S_i \Big|_{i=1}^{n}$, and the j^{th} member in $S_i \Big|_{i=1}^{n}$ is a vector $\left[p_x^{(i,j)}, p_y^{(i,j)}, p_h^{(i,j)}\right]$, which represents the horizontal coordinate, vertical coordinate, and elevation value of the j^{th} point, respectively. The information of each point in the DSM to be output is represented by a set S, whose j^{th} member is a vector $\left[p_x^j, p_y^j\right]$, representing the horizontal and vertical coordinates of the point, and whose final result requires the output of the corresponding p_h^j. Then, the elevation of any point of the DSM of the fused building is obtained.

The elevation values of the fusion results of n sets of DSMs need to satisfy two points, one is that the difference between the elevation of the corresponding approximate homonymous points before fusion is small enough, and the other is that they are as consistent as possible with the potential roof geometry of the building based on the spline function. To address the first point, the confidence level of each point in the DSM should be determined first. Similar to the method of confidence definition in the fusion method of surface DSM, if the stereo solutions of the near homonymous points of different DSMs are correct, they should have similar elevations. According to this assumption, this paper used the similarity of elevation of the near homonymous points as the confidence degree of the corresponding points. For the second point, the mesh partitioning of the spline function needs to be defined. To fit buildings with different complex structures, the strategy of iterative adaptive meshing was utilized to construct the proposed DSM fusion algorithm. Considering that most of the buildings are flat and pitched roofs, the grid is initialized as a low-scale grid. If the elevation deviation of the obtained fusion result is less than a given threshold, the current grid division is valid; otherwise, the grid scale is increased, and

recursive operations are performed. Based on the above two objectives, the optimization problem of n DSMs fusion can be constructed as shown in Equation (5).

$$\begin{aligned}
&\min_{P_h,\boldsymbol{a}_0,\boldsymbol{a}_1,\boldsymbol{b}_1} F_{fusion}(P_h,\boldsymbol{a}_0,\boldsymbol{a}_1,\boldsymbol{b}_1) = \\
&\sum_i\sum_j c_j\|p_h^j - p_h^{(i,Near_i(j))}\|_2^2 + \delta_1 \times \sum_j \sum_{[p_x^i,p_y^i]\in S_j} \left\| \left[a_0^j, a_1^j, b_1^j, -1\right] \times \left[1, p_x^i, p_y^i, p_h^i\right]^T \right\|_2^2 \\
&s.t. \left[a_0^j, a_1^j, b_1^j\right] \times \left[1, n_x^i, n_y^i\right]^T = \left[a_0^k, a_1^k, b_1^k\right] \times \left[1, n_x^i, n_y^i\right]^T \forall j \in node(i), k \in node(i) \\
&\left[a_1^j, b_1^j\right] \times \left[l_x^i, l_y^i\right]^T = \left[a_1^k, b_1^k\right] \times \left[l_x^i, l_y^i\right]^T \forall j \in line(i), k \in line(i)
\end{aligned} \tag{5}$$

where $\left[a_0^j, a_1^j, b_1^j\right]$, $\left[1, n_x^i, n_y^i\right]$, $\left[l_x^i, l_y^i\right]$, $node(i)$ and $line(i)$ are the parameters of the j^{th} spline function, the gridline intersection coordinates, the gridline slope parameter, the spline label adjacent to node i, and the spline label adjacent to the gridline, respectively. The left half of the objective function represents the elevation residuals of the DSM before and after minimizing the fusion, and the right half represents the elevation residuals of the DSM and the spline function after minimizing the fusion. Constraints are used to control the continuity of the spline function at the grid lines. The above equation is a nonconvex optimization problem, and in order to solve the problem, the idea of alternating iterations is used to fix the variable $[\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{b}_1]^T$ when solving for the variable P_h. At this time, the constraints do not contain variables, and the optimization problem is an unconstrained convex optimization problem, which can be obtained from the equation about P_h by setting the gradient to 0 by the objective function, as shown in Equation (6).

$$\frac{1}{N}\sum_i diag(C)\left[P_h - P_h^i\right] - \delta_1 \times \sum_j \sum_{[p_x^i,p_y^i]\in S_j} \left[a_0^j, a_1^j, b_1^j, -1\right] \times \left[1, p_x^i, p_y^i, p_h^i\right]^T = 0 \tag{6}$$

Therefore, the optimal solution P_h can be expressed as Equation (7).

$$P_h^* = \left(\frac{1}{N}\sum_i diag(C) - diag(-1_N)\right)^{-1} \left(\delta_1 \times \sum_j \sum_{[p_x^i,p_y^i]\in S_j} \left[a_0^j, a_1^j, b_1^j\right] \times \left[1, p_x^i, p_y^i\right]^T + \frac{1}{N}\sum_i diag(C)\left[P_h^i\right]\right) \tag{7}$$

Subsequently, the obtained P_h^* is fixed and solved for the variable $[\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{b}_1]^T$. At this point, the optimization problem is an equation-constrained convex optimization problem, and the original problem is first transformed into a Lagrangian function to remove the equation constraint, as shown in Equation (8).

$$\begin{aligned}
&L(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{b}_1, \gamma) = \\
&\sum_i\sum_j c_j\|p_h^j - p_h^{(i,Near_i(j))}\|_2^2 + \delta_1 \times \sum_j \sum_{[p_x^i,p_y^i]\in S_j} \left\| \left[a_0^j, a_1^j, b_1^j, -1\right] \times \left[1, p_x^i, p_y^i, p_h^i\right]^T \right\|_2^2 + \gamma A[\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{b}_1]^T
\end{aligned} \tag{8}$$

where $A[\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{b}_1]^T = 0$ is the collapsed equation constraint, γ is the dual variable. Further, the Lagrangian dual function is shown in Equation (9).

$$\begin{aligned}
&G(\gamma) = \min_{P_h} \sum_i\sum_j c_j\|p_h^j - p_h^{(i,Near_i(j))}\|_2^2 + \\
&\delta_1 \times \sum_j \sum_{[p_x^i,p_y^i]\in S_j} \left\| \left[a_0^j, a_1^j, b_1^j, -1\right] \times \left[1, p_x^i, p_y^i, p_h^i\right]^T \right\|_2^2 + \gamma A[\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{b}_1]^T
\end{aligned} \tag{9}$$

According to the slater condition, when the objective function is quadratic programming and the constraints are linear constraints, the dual gap is 0, and the optimal solution of the original problem coincides with the optimal solution of the dual problem. Since the KKT condition is a sufficient condition for obtaining the optimal solution of the Lagrangian

dual problem, the optimal solution of the dual problem can be solved through the KKT condition, as shown in Equation (10).

$$\begin{aligned}
&\frac{\partial L}{\partial\left[\boldsymbol{a}_0^*,\boldsymbol{a}_1^*,\boldsymbol{b}_1^*\right]}=0\Rightarrow\\
&\delta_1\times\sum_j\sum_{[p_x^i,p_y^i]\in S_j}\left[1,p_x^i,p_y^i\right]^T\times\left[1,p_x^i,p_y^i,p_h^i\right]\times\left[a_0^{j*},a_1^{j*},b_1^{j*},-1\right]^T+\gamma^TA=0\\
&A\left[\boldsymbol{a}_0^*,\boldsymbol{a}_1^*,\boldsymbol{b}_1^*\right]^T=0
\end{aligned}\tag{10}$$

To solve for the optimal $\left[\boldsymbol{a}_0^*,\boldsymbol{a}_1^*,\boldsymbol{b}_1^*\right]$, the above equation is replaced by $\left[\boldsymbol{a}_0^*,\boldsymbol{a}_1^*,\boldsymbol{b}_1^*\right]=[\boldsymbol{a}_0,\boldsymbol{a}_1,\boldsymbol{b}_1]+[\Delta\boldsymbol{a}_0,\Delta\boldsymbol{a}_1,\Delta\boldsymbol{b}_1]$ and rewritten in matrix equation form (see Equation (11)) to further solve for $[\Delta\boldsymbol{a}_0,\Delta\boldsymbol{a}_1,\Delta\boldsymbol{b}_1]$.

$$\begin{aligned}
&\begin{bmatrix}H,A^T\\A,0\end{bmatrix}[\Delta\boldsymbol{a}_0,\Delta\boldsymbol{a}_1,\Delta\boldsymbol{b}_1,\gamma]^T=\begin{bmatrix}-G\\0\end{bmatrix}\\
&G=\delta_1\times\sum_j\sum_{[p_x^i,p_y^i]\in S_j}\left[1,p_x^i,p_y^i\right]^T\times\left[1,p_x^i,p_y^i,p_h^i\right]\times\left[a_0^j,a_1^j,b_1^j,-1\right]^T\\
&H=\delta_1\times\sum_j\sum_{[p_x^i,p_y^i]\in S_j}\left[1,p_x^i,p_y^i\right]^T\times\left[1,p_x^i,p_y^i\right]
\end{aligned}\tag{11}$$

Solve this matrix equation to obtain $[\Delta\boldsymbol{a}_0,\Delta\boldsymbol{a}_1,\Delta\boldsymbol{b}_1]$, and update the optimization variables according to $\left[\boldsymbol{a}_0^*,\boldsymbol{a}_1^*,\boldsymbol{b}_1^*\right]=[\boldsymbol{a}_0,\boldsymbol{a}_1,\boldsymbol{b}_1]+[\Delta\boldsymbol{a}_0,\Delta\boldsymbol{a}_1,\Delta\boldsymbol{b}_1]$, alternately optimize P_h and $[\boldsymbol{a}_0,\boldsymbol{a}_1,\boldsymbol{b}_1]^T$, until $\left|F_{fusion}\left(P_h^n,\boldsymbol{a}_0^n,\boldsymbol{a}_1^n,\boldsymbol{b}_1^n\right)-F_{fusion}\left(P_h^{n-1},\boldsymbol{a}_0^{n-1},\boldsymbol{a}_1^{n-1},\boldsymbol{b}_1^{n-1}\right)\right|\le T_{threshold}$ is satisfied or the number of iterations exceeds a given threshold. When the iterations converge, the average elevation deviation between the fused DSM and the spline function is compared, that is, $\frac{1}{N_{DSM}}\sum_j\sum_{[p_x^i,p_y^i]\in S_j}\left\|\left[a_0^j,a_1^j,b_1^j,-1\right]\times\left[1,p_x^i,p_y^i,p_h^i\right]^T\right\|_2^2$, where N_{DSM} is the number of points in the fused DSM. If the average elevation deviation is less than the threshold, the current scale of meshing can effectively describe the building roof structure; otherwise, the scale of meshing is increased until the elevation deviation is less than the threshold or the number of iterations exceeds the limit. The DSM fusion method based on adaptive spline guided by building target structure characteristics can be summarized in Algorithm 1.

Algorithm 1. DSM fusion method based on adaptive spline guided by building target structure characteristics.

Input: n sets of DSMs $S_i\Big|_{i=1}^{n}$

Initialization: Initialize the mesh partitioning of the spline function, calculate the n sets of DSM weights C and the iteration counter $n_1=0,n_1=0$, initialize P_h^n and $\left[\boldsymbol{a}_0^n,\boldsymbol{a}_1^n,\boldsymbol{b}_1^n\right]^T$.

Step 1: $n_1=n_1+1$

Step 2: $n_2=n_2+1$

Step 3: Use alternate iteration method to optimize variables, fix $\left[\boldsymbol{a}_0^n,\boldsymbol{a}_1^n,\boldsymbol{b}_1^n\right]^T$ and P_h^n in turn, and use Equations (3) and (7) to calculate P_h^* and $[\Delta\boldsymbol{a}_0,\Delta\boldsymbol{a}_1,\Delta\boldsymbol{b}_1]$.

Step 4: Let $P_h^n=P_h^*$, $\left[\boldsymbol{a}_0^n,\boldsymbol{a}_1^n,\boldsymbol{b}_1^n\right]^T=[\boldsymbol{a}_0,\boldsymbol{a}_1,\boldsymbol{b}_1]^T+[\Delta\boldsymbol{a}_0,\Delta\boldsymbol{a}_1,\Delta\boldsymbol{b}_1]$.

Step 5: If $\left|F_{fusion}\left(P_h^n,\boldsymbol{a}_0^n,\boldsymbol{a}_1^n,\boldsymbol{b}_1^n\right)-F_{fusion}\left(P_h^{n-1},\boldsymbol{a}_0^{n-1},\boldsymbol{a}_1^{n-1},\boldsymbol{b}_1^{n-1}\right)\right|\le T_{threshold}$ or $n_2\ge n_{th}^2$ then execute Step 6; otherwise, skip to Step 2.

Step 6: If $\frac{1}{N_{DSM}}\sum_j\sum_{[p_x^i,p_y^i]\in S_j}\left\|\left[a_0^j,a_1^j,b_1^j,-1\right]\times\left[1,p_x^i,p_y^i,p_h^i\right]^T\right\|_2^2\le h_{th}$ or $n_1\ge n_{th}^1$ then the algorithm is executed; otherwise, increase the mesh density and skip to Step 1.

Output: DSM after fusion $S=\{P^{n_2}|1\le i\le N_{DSM}\}$.

3. Results

3.1. Experimental Results and Analysis of Simulated DSM Data

In order to verify the performance of the fusion algorithm under different roof shape building conditions, firstly, the target DSM with different roof shapes was generated as the truth map using computer simulation, and then Gaussian noise with different variances was generated and superimposed on the truth map DSM to generate two sets of observations, called DSM 1 and DSM 2. The two sets of noise-doped DSM 1 and DSM 2 were fused by executing the proposed DSM fusion algorithm, and the fusion results are shown in Figures 4–6.

It can be seen from the figure that after the noise is superimposed, the two sets of observed inputs DSM1 and DSM2 produce different degrees of difference from the truth map. When the noise variance is greater than 0.5, the roof structure of the building has undergone major deformation. By executing the proposed adaptive spline-based target characteristic-guided DSM fusion algorithm, the potential roof shapes in the DSMs to be fused are explored to the maximum extent to guide the generation of DSMs with high confidence. It can be seen from Figure 5d,g,j that the fused DSMs have the same roof geometry as the truth map, especially when the noise variance is 1. The proposed method still obtains excellent fusion results, and the comparison of elevation errors before and after fusion is shown in Tables 1–3.

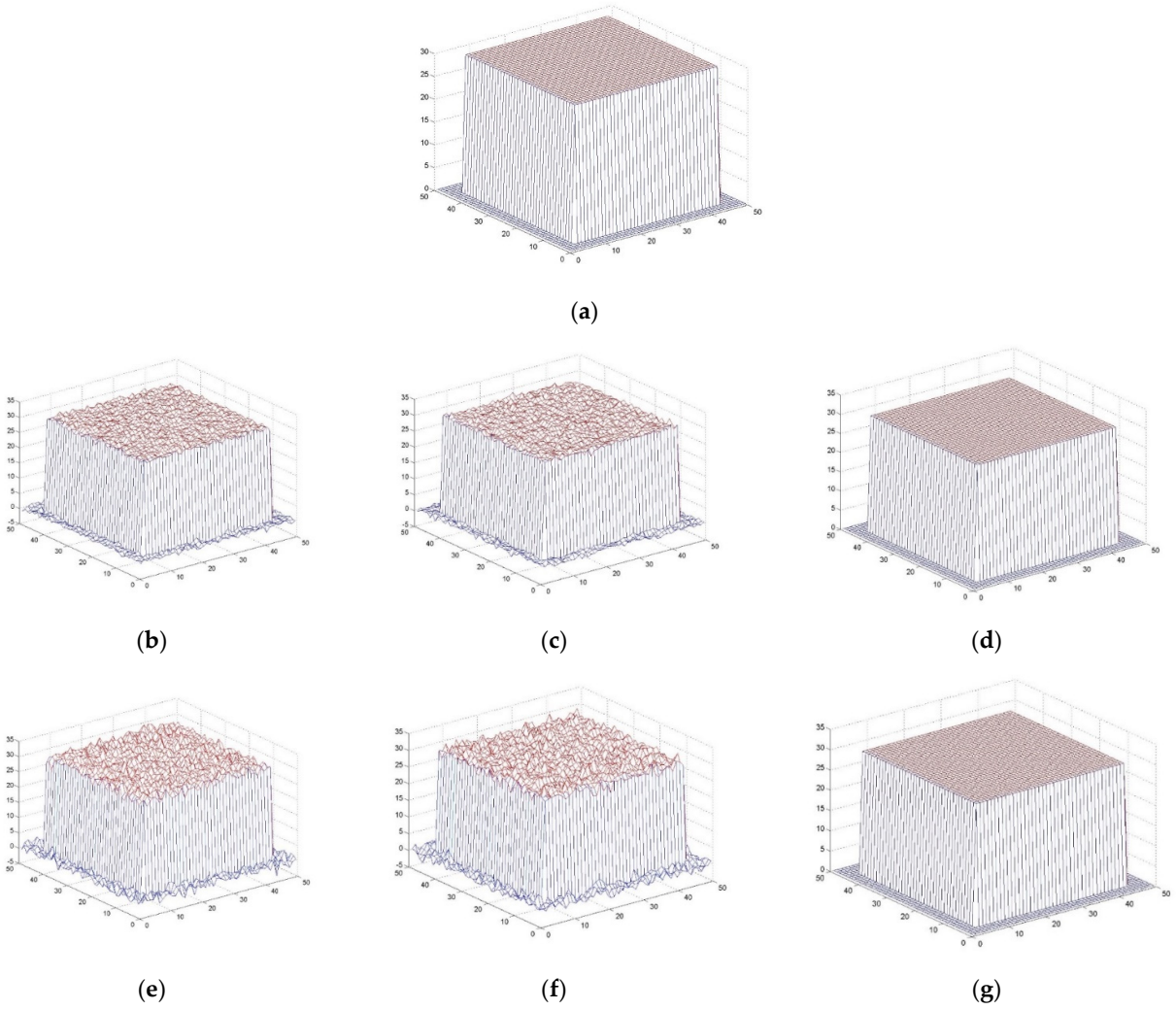


Figure 4. Comparison of the effect before and after the fusion of flat-roofed buildings with different variance Gaussian noise. (**a**) DSM of the truth map for flat-roofed building. (**b**) DSM 1 variance 0.5. (**c**) DSM 2 variance 0.5. (**d**) DSM after fusion variance 0.5. (**e**) DSM 1 variance 1. (**f**) DSM 2 variance 1. (**g**) DSM after fusion variance 1.

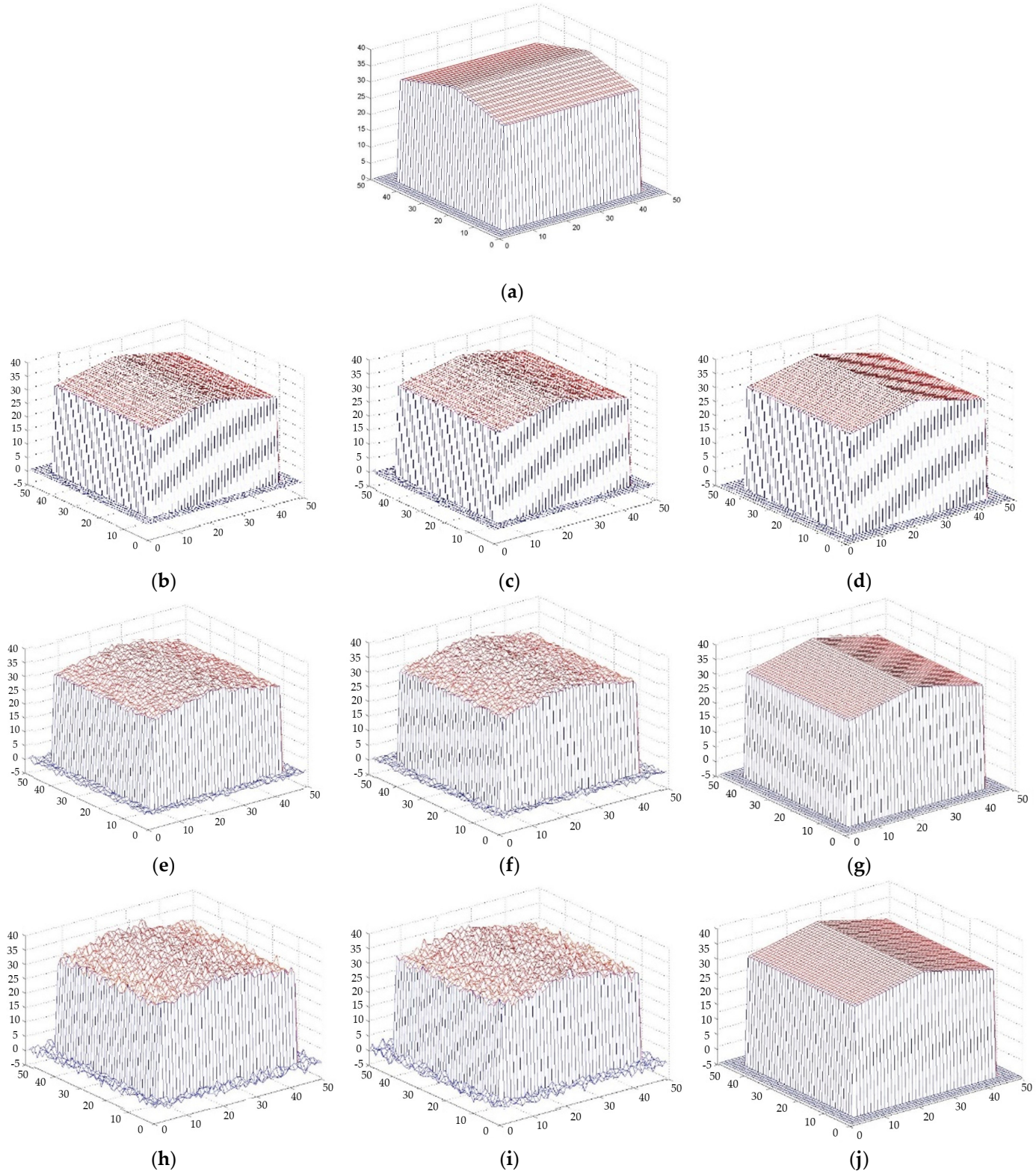


Figure 5. Comparison of the effect before and after the fusion of pitched-roofed buildings with different variance Gaussian noise. (**a**) DSM of the truth map for pitched-roofed building. (**b**) DSM 1 variance 0.1. (**c**) DSM 2 variance 0.1. (**d**) DSM after fusion variance 0.1. (**e**) DSM 1 variance 0.5. (**f**) DSM 2 variance 0.5. (**g**) DSM after fusion variance 0.5. (**h**) DSM 1 variance 1. (**i**) DSM 2 variance 1. (**j**) DSM after fusion variance 1.

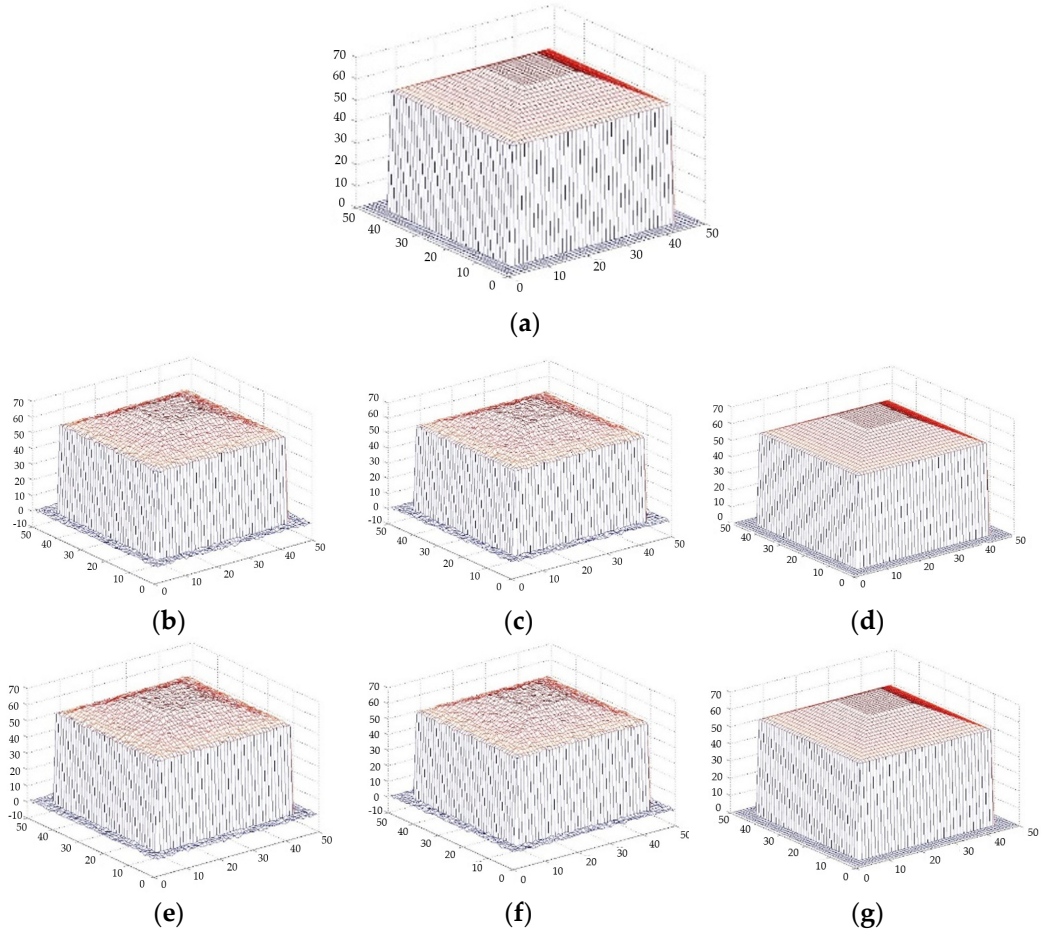


Figure 6. Comparison of the effect before and after the fusion of four-slope-roofed buildings with different variance Gaussian noise. (**a**) DSM of the truth map for four-slope-roofed building. (**b**) DSM 1 variance 0.5. (**c**) DSM 2 variance 0.5. (**d**) DSM after fusion variance 0.5. (**e**) DSM 1 variance 1. (**f**) DSM 2 variance 1. (**g**) DSM after fusion variance 1.

Table 1. Comparison of elevation errors before and after fusion of flat-roofed buildings.

	DSM 1	DSM 2	DSM after Fusion
Noise variance 0.5	0.3934	0.3976	0.0128
Noise variance 1	0.8120	0.7076	0.0135

Table 2. Comparison of elevation errors before and after fusion of pitched-roofed buildings.

	DSM 1	DSM 2	DSM after Fusion
Noise variance 0.1	0.0810	0.0801	0.0762
Noise variance 0.5	0.3942	0.4076	0.1266
Noise variance 1	0.8226	0.7983	0.1268

Table 3. Comparison of elevation errors before and after fusion of four-slope-roofed buildings.

	DSM 1	DSM 2	DSM after Fusion
Noise variance 0.5	0.3858	0.4095	0.0203
Noise variance 1	0.7961	0.7609	0.0320

The elevation error calculation is shown in Equation (12).

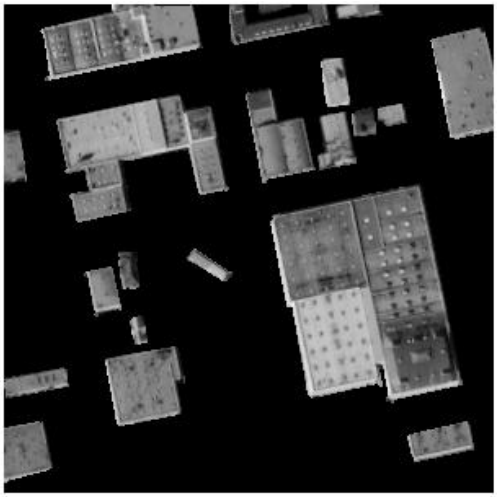
$$\sqrt{\frac{1}{m}\sum_{i=1}^{m}(h_i - \hat{h}_i)^2} \tag{12}$$

where h_i represents the true elevation value of the ground point, $\hat{h}_i$ represents the generated elevation value, and m represents the total number of ground points.

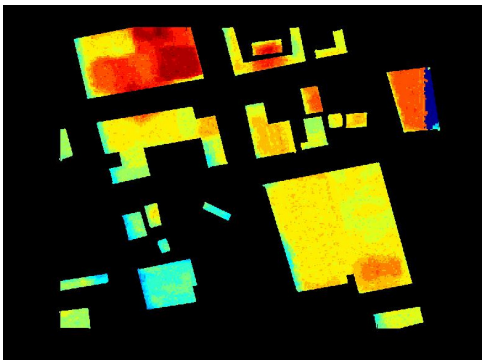
Through Tables 1–3, it can be found that the proposed DSM fusion algorithm can still significantly improve the DSM accuracy of the target under different roof shape building conditions, even in the case of large noise variance.

3.2. IKONOS Data Experiment Results

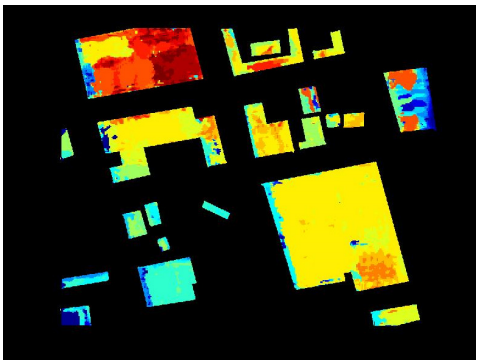
In this paper, two different stereo matching algorithms are used to generate two sets of DSMs for IKONOS data: Method 1 is a stereo matching algorithm based on the modified Census transform [29], Method 2 is a robust stereo matching algorithm using adaptive random walk with restart [30], and then the proposed DSM fusion algorithm is executed to obtain the final DSM. Combined with the truth map to compare and analyze the errors, the experimental results are shown in Figure 7, and the RMSEs before and after DSM fusion are shown in Table 4. The experimental results showed that the proposed method can effectively constrain and improve the DSM of the building, and the integrity of the overall construction of the target 3D model structure was significantly improved.



(**a**)

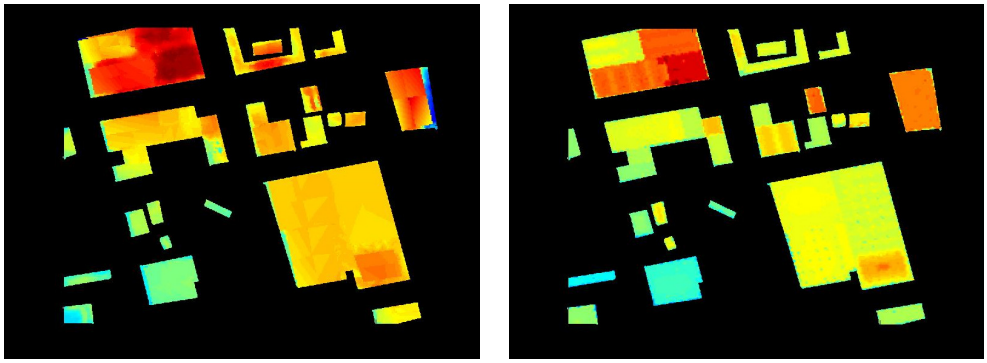


(**b**)



(**c**)

Figure 7. *Cont*.



(**d**) (**e**)

Figure 7. Comparison of DSM of buildings before and after fusion. (**a**) Original image. (**b**) DSM obtained by Method 1. (**c**) DSM obtained by Method 2. (**d**) DSM after fusion. (**e**) Truth map.

Table 4. Comparison of RMSE before and after fusion.

Method	RMSE
Method 1	2.24
Method 2	2.87
Fusion	2.16

4. Discussion

In this paper, we studied the DSM fusion method for building targets. In order to obtain better DSM results for buildings, a novel building DSM fusion method based on adaptive splines and target characteristic guidance was proposed. The improved DSM was obtained by assigning different DSM weights to integrate the advantages of multiple DSMs. Based on the existing prior knowledge, the better the DSM effect, the higher the weight assigned; otherwise, it is treated as equal weight.

This paper used adaptive spline functions to fit the roof structure of different buildings. Since the spline function is a piecewise function, it can be divided into an appropriate interval to accurately describe the specific building roof structure. At the same time, through rigorous mathematical proofs, we concluded that the denser the grid, the more accurately the spline function can describe the complex roof shape of a building. Therefore, this paper constructed a set of adaptive-meshing low-order spline functions to describe the characteristics of the target structure. The adaptive spline method can be applied in many aspects [31–33].

The part of the local deformation of the building in the DSM generated by the inaccurate elevation will also be fitted by the spline function. The DSM obtained directly is not accurate enough, and many existing studies are based on the constraint DSM [1,3,5,27,34]. In order to maintain the building roof structure as much as possible and avoid the introduction of inaccurate elevation information in the DSM, a DSM fusion method based on adaptive spline guided by building geometric structure characteristics was constructed to obtain a more accurate DSM of the building by fusing multiple DSMs. The acquired DSMs can come from different sources, either from LIDAR or from stereo reconstruction [29,30]. We took the fusion of two DSMs as an example, used simulation to generate DSMs with different roof shapes, generated two sets of observation data by superimposing Gaussian noise with different variances in the generated DSMs, then performed fusion experiments using the proposed method. At the same time, we also used IKONOS data with different stereo matching algorithms to generate the corresponding DSMs and perform fusion. The greater the number of DSMs, the better the fusion effect will be. However, as the number of DSMs increases, the corresponding computational consumption will also increase. In order to balance the number of DSMs with computational consumption, the number of DSMs is not as large as possible.

In summary, the proposed method can effectively preserve the edges and the integrity of the DSM of buildings. The experimental results show that the proposed method is an effective and efficient DSM constraint method for buildings.

5. Conclusions

This paper proposed an adaptive spline and target characteristic-guided DSM fusion update method for buildings and conducted DSM fusion experiments using typical urban area images of different scenes. The experimental results showed that the proposed method can effectively constrain and improve the DSM of buildings, and the integrity of the overall construction of the target 3D model structure was significantly improved, which is an effective and efficient DSM constraint method for buildings.

Author Contributions: Conceptualization, methodology, visualization, and writing—original draft preparation, J.L.; data curation, formal analysis, supervision, project administration, funding acquisition, and writing—review and editing, H.C.; investigation, software, resources, and validation, S.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, Grant Number: 61771170.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ruchay, A.; Dorofeev, K.; Kolpakov, V. Fusion of information from multiple kinect sensors for 3D object reconstruction. *Comput. Opt.* **2018**, *42*, 898–903. [CrossRef]
2. Zhao, L.; Liu, Y.; Men, C.; Men, Y. Double Propagation Stereo Matching for Urban 3-D Reconstruction from Satellite Imagery. *IEEE Trans. Geosci. Remote Sens.* **2021**, 1–17. [CrossRef]
3. Partovi, T.; Fraundorfer, F.; Bahmanyar, R.; Huang, H.; Reinartz, P. Automatic 3-D building model reconstruction from Very High Resolution stereo satellite imagery. *Remote Sens.* **2019**, *11*, 1660. [CrossRef]
4. Jiao, N.; Wang, F.; You, H. A New RD-RFM Stereo Geolocation Model for 3D Geo-Information Reconstruction of SAR-Optical Satellite Image Pairs. *IEEE Access* **2020**, *8*, 94654–94664. [CrossRef]
5. Li, Y.; Wu, B. Relation-Constrained 3D Reconstruction of Buildings in Metropolitan Areas from Photogrammetric Point Clouds. *Remote Sens.* **2021**, *13*, 129. [CrossRef]
6. Rupnik, E.; Pierrot-Deseilligny, M.; Delorme, A. 3D reconstruction from multi-view VHR-satellite images in MicMac. *ISPRS J. Photogramm. Remote Sens.* **2018**, *139*, 201–211. [CrossRef]
7. Oude Elberink, S.; Vosselman, G. 3D information extraction from laser point clouds covering complex road junctions. *Photogramm. Rec.* **2009**, *24*, 23–36. [CrossRef]
8. Zomer, R.; Ustin, S.; Ives, J. Using satellite Remote Sensing for DEM extraction in complex mountainous terrain: Landscape analysis of the Makalu Barun National Park of eastern Nepal. *Int. J. Remote Sens.* **2002**, *23*, 125–143. [CrossRef]
9. Arefi, H.; Reinartz, P. Building reconstruction using dsm and orthorectified images. *Remote Sens.* **2013**, *5*, 1681–1703. [CrossRef]
10. Xu, B.; Jiang, W.; Shan, J.; Zhang, J.; Li, L. Investigation on the Weighted RANSAC Approaches for Building Roof Plane Segmentation from LiDAR Point Clouds. *Remote Sens.* **2016**, *8*, 5. [CrossRef]
11. Wang, H.; Zhang, W.; Chen, Y.; Chen, M.; Yan, K. Semantic decomposition and reconstruction of compound buildings with symmetric roofs from LiDAR data and aerial imagery. *Remote Sens.* **2015**, *7*, 13945–13974. [CrossRef]
12. Karantzalos, K.; Paragios, N. Large-scale building reconstruction through information fusion and 3-D priors. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2283–2296. [CrossRef]
13. Oude Elberink, S.; Vosselman, G. Building reconstruction by target based graph matching on incomplete laser data: Analysis and limitations. *Sensors* **2009**, *9*, 6101–6118. [CrossRef] [PubMed]
14. Brenner, C. Constraints for modelling complex objects. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2005**, *36*, 49–54.
15. Henn, A.; Gröger, G.; Stroh, V.; Plümer, L. Model driven reconstruction of roofs from sparse lidar point clouds. *ISPRS J. Photogramm. Remote Sens.* **2013**, *76*, 17–29. [CrossRef]
16. Lafarge, F.; Descombes, X.; Zerubia, J.; Pierrot-Deseilligny, M. Structural approach for building reconstruction from a single DSM. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 135–147. [CrossRef]
17. Huang, H.; Brenner, C.; Sester, M. A generative statistical approach to automatic 3D building roof reconstruction from laser scanning data. *ISPRS J. Photogramm. Remote Sens.* **2013**, *79*, 29–43. [CrossRef]
18. Habib, A.F.; Zhai, R.; Kim, C. Generation of complex polyhedral building models by integrating stereo-aerial imagery and lidar data. *Photogramm. Eng. Remote Sens.* **2010**, *76*, 609–623. [CrossRef]

19. Awrangjeb, M.; Fraser, C.S. Automatic segmentation of raw Lidar data for extraction of building roofs. *Remote Sens.* **2014**, *6*, 3716–3751. [CrossRef]
20. Chen, D.; Wang, R.; Peethambaran, J. Topologically Aware Building Rooftop Reconstruction From Airborne Laser Scanning Point Clouds. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 7032–7052. [CrossRef]
21. Verma, V.; Kumar, R.; Hsu, S. 3D Building Detection and Modeling From Aerial Lidar Data. In Proceedings of the IEEE Computer Society Conference on the Computer Vision and Pattern Recognition, New York, NY, USA, 17–22 June 2006; Volume 2, pp. 2213–2220.
22. Chauve, A.-L.; Labatut, P.; Pons, J.-P. Robust piecewise-planar 3D reconstruction and completion from large-scale unstructured point data. In Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, CA, USA, 13–18 June 2010; pp. 1261–1268.
23. Sampath, A.; Shan, J. Segmentation and reconstruction of polyhedral building roofs from aerial LIDAR point clouds. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 1554–1567. [CrossRef]
24. Kim, K.; Shan, J. Building roof modeling from airborne laser scanning data based on level set approach. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 484–497. [CrossRef]
25. Tian, Y.; Gerke, M.; Vosselman, G.; Zhu, Q. Knowledge-based building reconstruction from terrestrial video sequences. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 395–408. [CrossRef]
26. Fan, H.; Yao, W.; Fu, Q. Segmentation of Sloped Roofs from Airborne LiDAR point Clouds Using Ridge-based Hierarchical Decomposition. *Remote Sens.* **2014**, *6*, 3284–3301. [CrossRef]
27. Xiong, B.; Elberink, S.O.; Vosselman, G. A graph edit dictionary for correcting errors in roof topology graphs reconstructed from point clouds. *ISPRS J. Photogramm. Remote Sens.* **2014**, *93*, 227–242. [CrossRef]
28. Yin, X.; Hao, X.; Gao, T.; Chen, H.; Chen, W. Building Detection based on Rectangle Approximation and Region Growing. In Proceedings of the 2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa Village, HI, USA, 16–26 July 2020.
29. Ma, L.; Li, J.; Ma, J.; Zhang, H. A Modified Census Transform Based on the Neighborhood Information for Stereo Matching Algorithm. In Proceedings of the 2013 Seventh International Conference on Image and Graphics, Qingdao, China, 26–28 July 2013.
30. Lee, S.; Jin, H.L.; Lim, J.; Suh, I.H. Robust stereo matching using adaptive random walk with restart algorithm. *Image Vis. Comput.* **2015**, *37*, 1–11. [CrossRef]
31. Liu, Z.; Guo, W. Data Driven Adaptive Spline Smoothing. *Stat. Sin.* **2010**, *20*, 1143–1163.
32. Wang, W.; Zhao, H.; Zeng, X.; Doğançay, K. Steady-state performance analysis of nonlinear spline adaptive filter under maximum correntropy criterion. *IEEE Trans. Circuits Syst. II Exp. Briefs* **2020**, *67*, 1154–1158. [CrossRef]
33. Roy, S.S.; Roy, R.; Balas, V.E. Estimating heating load in buildings using multivariate adaptive regression splines, extreme learning machine, a hybrid model of MARS and ELM. *Renew. Sustain. Energy Rev.* **2018**, *82*, 4256–4268.
34. Xu, B.; Jiang, W.; Li, L. HRTT: A hierarchical roof topology structure for robust building roof reconstruction from point clouds. *Remote Sens.* **2017**, *9*, 354. [CrossRef]