WILEY | Hindawi

Retraction

Retracted: Application of Remote Sensing Image Data Scene Generation Method in Smart City

Complexity

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Xu, "Application of Remote Sensing Image Data Scene Generation Method in Smart City," *Complexity*, vol. 2021, Article ID 6653841, 13 pages, 2021.

WILEY | Hindawi

Research Article

Application of Remote Sensing Image Data Scene Generation Method in Smart City

Yuanjin Xu [ORCID]

Institute of Mathematical Geology and Remote Sensing Geology, School of Earth Resources, China University of Geosciences, 388 Lumo Road, Wuhan 430074, China

Correspondence should be addressed to Yuanjin Xu; xuyj@cug.edu.cn

Remote sensing image simulation is a very effective method to verify the feasibility of sensor devices for ground observation. The key to remote sensing image application is that simultaneous interpreting of remote sensing images can make use of the different characteristics of different data, eliminate the redundancy and contradiction between different sensors, and improve the timeliness and reliability of remote sensing information extraction. The hotspots and difficulties in this direction are based on remote sensing image simulation of 3D scenes on the ground. Therefore, constructing the 3D scene model on the ground rapidly and accurately is the focus of current research. Because different scenes have different radiation characteristics, therefore, when using MATLAB to write a program generated by 3D scenes, 3D scenes must be saved as different text files according to different scene types, and then extension program of the scene is written to solve the defect that the calculation efficiency is not ideal due to the huge amount of data. This paper uses POV ray photon reverse tracking software to simulate the imaging process of remote sensing sensors, coordinate transformation is used to convert a triangle text file to POV ray readable information and input the RGB value of the base color based on the colorimetry principle, and the final 3D scene is visualized. This paper analyzes the thermal radiation characteristics of the scene and proves the rationality of the scene simulation. The experimental results show that introducing the chroma in the visualization of the scene model makes the whole scene have not only fidelity, but also radiation characteristics in shape and color. This is indispensable in existing 3D modeling and visualization studies. Compared with the complex radiation transmission method, using the multiple angle two-dimensional image generated by POV rays to analyze the radiation characteristics of the scene, the result is intuitive and easy to understand.

1. Introduction

With the development of science and technology, digital cities have received more and more attention [1]. The concept of digital city originates from the strategic concept of digital earth, also known as network city or smart city, or more precisely information city. It refers to the comprehensive use of computer tools (GIS, remote sensing, telemetry, network, multimedia, and virtual simulation technology). The use of digital technology to collect and process the city's infrastructure and functional mechanisms to enable it to have digital functions, which is conducive to optimizing and improving the city's ecological environment and resources, economy, population, and other complex fields, effectively predicts the future of the city. [2, 3]. The

essence or core of a digital city is the fusion of massive urban spatial data with three-dimensional urban geographic information systems and time series urban geographic information systems [4]. The outstanding feature of digital cities is the ability of applying digital information to grasp the changing process of urban regional structure in time and space. The application research of 3D urban geographic information system and time series urban geographic information system will be an important part of recent digital city theory research. Digital city construction will provide an information security system for the city's sustainable development strategy, meet the government's decision-making, macrocontrol, scientific and technological innovation, natural resources and environmental monitoring, intelligent transportation and urban

management, and various social welfare undertakings, and further provide solutions for the sustainable development of cities. [5, 6]. To build a digital city, we must first apply high-tech means such as computer technology to model the urban environment [7].

Remote sensing image fusion is a process of comprehensive processing the image data obtained by multiple remote sensing sensors or the same kind of sensor for the same target at different times. The image is processed by using certain rules or algorithms, and the useful information contained in the image is fused into a new image. The image contains more accurate and abundant information than any single image, in order to achieve a comprehensive description of the target and ground objects.

The problem of three-dimensional reconstruction of urban buildings has been studied by experts and scholars from various countries for many years and has achieved a series of results. The most representative ones are Google Earth and Microsoft Virtual Earth, which use satellite remote sensing images to generate virtual ground scenes, which have been successfully commercialized on the Internet [8, 9]. Image-based three-dimensional reconstruction of urban buildings is mainly divided into three categories according to the different data sources used: (1) based on remote sensing images, this method uses three-dimensional reconstruction of urban buildings using approximately vertical satellite remote sensing images or aerial images [10]. According to the characteristics of remote sensing imaging, based on the reconstruction of remote sensing images, the reconstruction space is large, and the roof information of the building can also be obtained and the accumulation of errors can be effectively reduced, but the reconstructed buildings have poor fidelity [11]. (2) Based on ground image, ground image-based reconstruction is a three-dimensional reconstruction of urban buildings using images acquired by various ground shooting techniques. According to the characteristics of ground imaging, the reconstruction of this method is better, and the wall texture of the building can be obtained, but the roof information of the building is not obtained, the reconstruction scale is small, and the error accumulation is large [12]. (3) Regarding combination of remote sensing image and ground image, there are advantages and disadvantages in remote sensing and ground-based imaging reconstruction. In fact, remote sensing image and ground image are two important complementary source data. Combining the two for reconstruction is expected to be obtained. This resulted in a reconstruction method combining remote sensing images with ground images. Generally, there are insufficient data acquisition costs, large data volume, complicated calculation, and low automation [13, 14].

Based on the second generation bending wave transform and Dempster-Shafer (DS) evidence theory, Huang C proposed a new remote sensing image fusion method. Huang C uses the bending wave transform to decompose the remote sensing image to obtain the coefficients and uses DS evidence theory to optimize the high coefficients [15, 16]. First, the high-resolution and multiple spectral remote sensing images are decomposed by bending wave transform

to obtain bending wave transform coefficients (coarse, detailed, and fine scale layers) of all layers. Second, the coarse scale layer uses the maximum fusion rule. The detailed scale layer is used by the weighted average fusion rule. The fine scale layer is optimized by DS evidence theory. Three features of the fine scale layer coefficients are obtained. These three characteristics are variance, information entropy, and energy. The use of these features is some parametric belief function and rationality function. The mass function is then combined and a new fusion factor is obtained. Finally, the scene image is obtained by inverse bending wave transform. Rhee et al. attempt to apply two types of image matching, object space based matching techniques and image space based matching techniques, and compare the performance of the two techniques [17, 18]. The object space based match used sets a list of candidate height values for a fixed horizontal position in the object space. For each height, its corresponding image points are calculated and similarity is measured by gray level correlation. The image space based matching used is a modified slack match. Rhee and Kim designed a global optimization scheme for finding the best pair (or group) to apply image matching, defining local matching regions in the image or object space, and merging local point clouds into global point clouds. For optimal pairing selection, the connection points between the images are extracted and a stereoscopic overlay network is defined by using the connection points to form a maximum spanning tree. Qin built the core technology and method related to 3D model reconstruction, focusing on matching of point cloud data registration to simplify denoising, 2D contour extraction, and finally achieving the high complexity of 3D geometric model of farmland site, using advanced 3D printing. Technology produces small 3D printed point cloud data [19, 20].

The innovations of this paper are as follows: (1) introducing the principle of colorimetry method into the visualization of the image simulation scene model and replacing the complex texture with color can reflect the spectral radiation characteristics of the object to a certain extent. Through investigation and research, it is found that the only one that can correspond to the spectral characteristics is its chroma characteristic. Therefore, the chroma is introduced into the scene model visualization so that the whole scene not only is in shape and color, but also has radiation characteristics. This is not available in existing 3D modeling and visualization studies. (2) The remote sensing imaging process can be simulated using the POV ray visual ray tracing software package. It has a convenient and fast programming language, high computational efficiency, and intuitive output. Compared with the complex radiation transfer equation, the multiple angle two-dimensional image generated by POV ray is used to analyze the radiation characteristics of the scene, and the result is intuitive and easy to understand. POV ray can simulate the remote sensing imaging process, mainly because it first defines the position of the light source (sun) and camera (sensor), zenith angle and azimuth, and also sets the camera's field of view, which is another 3D visualization software. And the higher computational efficiency is also relatively advanced in the

field of visualization. (3) Considering the differences of different scenes, the design of remote sensing images for different scenes is different, so that the experimental structure can reflect the diversity and rationality.

2. The Proposed Method

2.1. Preparation of Remote Sensing Images

2.1.1. Image Preprocessing. When the obtained source image is blurred, the contrast is not strong, and the noise interference is large; the corresponding method needs to be used for some processing, so that the subsequent work can be carried out more effectively. Common methods include image enhancement, filtering, histogram correction, and gradation transformation. However, when the quality of the source image is good, it is not necessary to perform these processes. Therefore, the preprocessing of the source image is an optional link, and the processing method is different for different images. The omnidirectional image is used in the experiment in this paper; because the camera has been calibrated before shooting, the lighting conditions are better at the time of shooting, so the obtained image quality is better and generally does not have to be preprocessed.

2.1.2. Registration. Registration is to find the mutual correspondence between the omnidirectional map and the remote sensing map and achieve the purpose of reconstruction service through information fusion. Information fusion is the foundation of all subsequent work, so registration is a core component of the entire reconstruction. Conventional methods generally solve the problem of registration between images from homologous images or from the same type of sensor. The omnidirectional and remote sensing images are images formed by heterogeneous sensors. The general reconstruction method of the source image requires registration of the source image, and more registration methods are available at present. However, the existing conventional methods cannot solve the registration problem of remote sensing maps well. The registration of remote sensing maps belongs to the specific key technical problems of 3D reconstruction of remote sensing images.

2.1.3. High Extraction. Height extraction is to obtain the height value of the target building in real space, which is an important information of the space structure of the building. The height of the building can be combined with the top view of the building available in the remote sensing map to obtain the approximate spatial structure of the building. Therefore, the height of the building plays a crucial role in the reconstruction of the shape and contour. There are two conventional solutions; one is direct measurement using instruments and equipment, such as laser range finder. The second is to use computer vision principles for estimation. Direct measurement with related instruments is costly. With computer vision estimation, accurate absolute heights are generally not obtained without accurate scale reference objects.

2.1.4. Shape Modeling. The goal of shape modeling is to get the outer shape of the entire building. It is also an important part of reconstruction. It is generally modeled based on certain assumptions or prior knowledge (such as a box in the shape of a box and a flat roof) using the obtained height and roof shape information obtained from the remote sensing map. The outer shape of the building is more complex and precise than the approximate spatial structure. The key to shape modeling is to extract the outline of the building. At present, the contour detection algorithms generally have shortcomings such as low detection rate and inability of fully automating. The method of semiautomatic human-computer interaction can be adopted; that is, the existing detection algorithm is combined with the manual correction method for detection. Due to limited time and energy, image-based 3D reconstruction mostly requires shape modeling, and shape modeling methods are almost universal.

2.2. Building Outline Segmentation and 3D Model Extraction. How to obtain the singular model of the building from the three-dimensional model of the scene generated by oblique photography is the goal of this paper. DOM can be seen from the high-resolution scene, which has obvious image difference between the building and other features. The image analysis method can be used to extract the outline of the building from the scene DOM and obtain the position information of the building outline, thereby realizing the positioning and segmentation of the building model in the three-dimensional scene model.

2.2.1. High-Resolution Image E Building Feature Analysis Method. In an image, the edge information is the most important and basic feature, and the edge feature is the most direct expression and embodiment of the image geometric information. The low-altitude drone oblique photography obtains a higher resolution of the image, and the texture of the scene DOM obtained after the correction is clear. The difference in shape, size, and texture patterns is the basic basis for distinguishing between different features. Through vision, the color of the scene DOM is very realistic, the texture information is very rich, the geometrical structure of the object is more refined, and the recognition of different target objects in the image is more accurate; from a local perspective, a single feature, especially the boundary between the edge of the building and its surrounding environment, is obvious, and the details inside the target object are richly expressed. These features are very advantageous for identifying and extracting the target individual of the building. However, because of the rich information contained in high-resolution images, the phenomenon of "homologous" and "homogeneous foreign matter" appears, resulting in increased noise interference.

In order to extract buildings (houses) in high-resolution images, the characteristics of the buildings are analyzed to establish a good basis for building identification. From spectral and texture characteristics, usually the gray distribution of buildings is relatively uniform. The gray value of the

top is higher than the gray value of other surrounding objects and the texture mode is relatively regular. Generally, the texture performance of buildings has the outline direction of the building that is approximately uniform or orthogonal. For the shape feature, usually the building has geometrically regular edges and corners, and the whole is presented as a regular polygon. From the spatial distribution characteristics, usually the roads in the urban area will be divided into several blocks like chessboards. The buildings are regularly distributed in the block, and the surrounding objects mainly include roads and tree vegetation, so the roads and buildings have strong spatial associations, and usually densely populated buildings will also be arranged regularly and have similar configurations. When considering the geometric features, spectral features, and spatial distribution characteristics of features in high-resolution images, the algorithm is not mature enough and complicated, having low efficiency, which needs further study. According to the research objectives and needs of this paper, the rough positioning of the target of the suspected building does not require accurate and complete extraction of the building. Therefore, this paper only analyzes the geometric features of the building from the geometric shape. In the scene DOM, the geometry of buildings and other features is significantly different. From a straight-down perspective, no matter which building is an individual wrapped by its outer contour, in the high-resolution scene DOM, the outer contour of the building is a connected area with a certain length and area. As a basis for identifying buildings, the internal structure of vegetation such as trees is disorderly, and the boundaries of the whole forest are not clear enough and there are no rules: there are often no clear and regular boundaries on both sides of the road, and the roads are lacking in certain areas. From the perspective of individual buildings, there are also separate topological relationships between different buildings. In the scene DOM, the target features that can be approximated as connected areas tend not to have only buildings, but the buildings also contain other distinguishable geometric features, and the more prominent common features are the length of the outer contour of the building. The outer contour of the building must contain at least a certain number of linear features, according to which the connected areas surrounded by the outline of the building can be further identified and screened.

Therefore, by analyzing the features of the building on the image, the difference and segmentation between the building and other features and buildings can be realized, and the outline of a single suspected building can be extracted and used as a basis for the three-dimensional scene model. A rough monomer model was extracted.

2.2.2. Building Edge Feature Detection Method Based on Canny Operator. Edge detection is to obtain information about shape and reflection or transmittance in an image. It is a basic step in image processing, analysis, understanding, pattern recognition, and computer and human vision, being

a very important technology. There are many edge detection methods, such as Roberts operator, Pruitt operator, Sobel operator, and Laplace operator. These are the operators that detect features through local window and are sensitive to noise. Canny proposed the best edge detection operator Canny operator. The operator determines the edge pixels by the maximum value of the image signal function, and the detection performance is good, which has been widely used. Therefore, this paper uses Canny operator to perform edge detection on scene digital image. The scene DOM is a true color image, which needs to be grayed out. The color image can be converted into a grayscale image by using the following formula:

$$P(x, y) = 0.3^*R + 0.59^*, G + 0.11^*B, \tag{1}$$

where $P(x, y)$ is the gray value of the pixel at the (x, y) coordinate and R, G, and B are the values in the red, green, and blue primary color channels in the pixel at the (x, y) coordinate, respectively. There are four main steps in detecting the edge features of grayscale images using the Canny operator:

(1) Eliminate Noise. The differential algorithm is highly sensitive to noise, and the Gaussian smoothing filter is used to convolve the image before edge detection to reduce noise interference. The first following formula is a two-dimensional Gaussian function. The principle of Gaussian smoothing is the discrete Gaussian function. The Gaussian function value on the discrete point is used as the weight. For each pixel in the gray image, it is within the window neighborhood of a certain size. Considering pixel weighting to eliminate Gaussian noise, the second following formula is a discrete Gaussian function weight window template with a window size of 5×5 pixels, the third following formul is a convolution formula for Gaussian filtering of image J, and g is the result of convolution:

$$f(x, y) = \frac{1}{2\pi\sigma^2}e^{((x^2+y^2)/2\sigma^2)}, \tag{2}$$

$$K = \frac{1}{159}\begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix}, \tag{3}$$

$$g = I^*K. \tag{4}$$

(2) Calculate the Image Gradient Amplitude Value and the Gradient Direction. The first-order finite difference is used to approximate the gray gradient of the image. In the Canny operator, the first following formula is used to convolve in x and y directions of the image. As shown in the second and third following formulas, the Sobel template is shown,

wherein S_x is a convolution template in the horizontal direction, and S_y is a convolution template in the vertical direction:

$$G_x = g^* S,$$
$$G_y = g^* S_y, \tag{5}$$

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \tag{6}$$

$$S_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}. \tag{7}$$

According to the following formulas, the gradient magnitude value G and the direction θ of the image can be, respectively, calculated and the gradient direction is approximated to be generally 0°, 45°, 90°, and 135°:

$$G = \sqrt{G_x^2 + G_y^2}, \tag{8}$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right). \tag{9}$$

(3) Nonmaximum Suppression. The pixel corresponding to the local maximum point is found and retained or marked as an edge pixel, and the gray value of the pixel of the nonmaximum point is suppressed and set as the background. This step is mainly to discriminate and remove nonedge pixels, leaving only the candidate image edges.

$$A = \sum \left[\text{Si}^* \log\left(\frac{\text{Si}}{\text{Qi}}\right) \right]. \tag{10}$$

The image edge is solved using a hysteresis threshold algorithm. The method of lag threshold is to use two thresholds (high and low). The following discriminant basis is used when determining the true edge and removing the false edge: if the gradient amplitude value of a pixel is greater than the high threshold, it is determined that the pixel is a true edge pixel and is retained; if the gradient amplitude value of a pixel is less than the low threshold, it is determined that the pixel is not a true edge pixel and is excluded; if the gradient amplitude value of a pixel is between the high and low thresholds, the pixel has only one gradient amplitude value. When pixels larger than the high threshold are connected, it is determined that the pixel is reserved for the real edge pixel. The scene digital orthophoto is rich in texture, and the geometric features of the features are complex. The edge obtained by the Canny operator for edge detection usually contains a lot of noise. Therefore, the DOM edge detection result needs to be Gaussian. In order to facilitate the identification of the building outline, it is necessary to highlight the contour edge features of the target such as a building from the numerous edge feature pieces of

information in the scene and further suppress the edge pixels whose edge features are not obvious into the background. Therefore, it is necessary to perform Gaussian smoothing on the scene edge detection result and then perform binary processing. Treated by formula (10), we get

$$g_{\text{new}}(x, y) = \begin{cases} 255 & (g_{\text{old}}(x, y) > T), \\ 0 & (g_{\text{old}}(x, y) \le T), \end{cases} \tag{11}$$

In the formula, $g_{\text{new}}(x, y)$ is the new gray value of the pixel at the (x, y) coordinate; $g_{\text{old}}(x, y)$ is the previous gray value of the pixel; T is the gray threshold set according to experimental experience and belongs to 0–255. If the gray value of a pixel is greater than the threshold, the gray value of the pixel is set to 255. If it is not greater than the threshold, the gray value is set to 0, and the background is suppressed, thereby obtaining a scene edge with a distinct target edge feature.

2.3. Construction of the Original Features of the Front Image. According to the previous preset, the output of the CNN convolutional layer is more than fully connected:

$$a_k = \sum_{y=1}^{H} \sum_{x=1}^{W} w_{x,y} b_{x,y} \cdot f_{x,y,k}. \tag{12}$$

Among them, $w_{x,y}$ is the response weight at the point (x, y) in the feature map, and $b_{x,y}$ is the depth weight of the point. For the response weight matrix $W \in R^{W \times H}$, we use the feature map f_k to construct

$$W = \sum_{k=1}^{N} f_k. \tag{13}$$

For the depth weight, we use the depth information of the image to assign the weight. We first scale the depth map of the input image to $W \times H$; then,

$$d_{x,y} = \left(\frac{d_{\max} - d(x, y)}{d_{\max} - d_{\min}}\right) + \gamma. \tag{14}$$

Among them, $d_{\max}$ is the maximum depth, $d_{\min}$ is the minimum depth, and $d(x, y)$ is the depth information at (x, y). γ is a very small amount to ensure that the depth estimation process of the monocular image is very far away (the misjudgment). The weight value tilts the image feature to the close range.

For the obtained weight matrix, we use the $L2$ norm to normalize; namely,

$$V = \frac{V}{\|V\|_2}. \tag{15}$$

We perform the above sum pooling calculation on all N feature maps output by the convolutional layer L to obtain an N-dimensional feature vector φ of the convolutional layer, and use the same-dimensional PCA whitening, and then perform the whitening features obtained. The $L2$ norm is normalized, and finally N-dimensional image features are obtained.

For comparison, we extract the SPOC algorithm as follows:

$$a_k = \sum_{y=1}^{H} \sum_{x=1}^{W} c_{x,y} f_{x,y,k}. \tag{16}$$

Among them, $c_{x,y}$ is the Gaussian center prior, and its weights are set as follows:

$$c_{x,y} = \exp\left(-\frac{(y - H/2)^2 + (x - W/2)^2}{2\partial^2}\right). \tag{17}$$

Among them, ∂ is the distribution covariance, which is set to one-third of the length of the feature map center from the nearest boundary. It can be seen that the SPOC algorithm adds a Gaussian center prior on the basis of sum pooling and does not effectively reflect the icon in the image. Objects cannot reflect the characteristics of close-up shots.

3. Experiments

3.1. Data Acquisition. Remote sensing is a means of collecting electromagnetic fields, force fields. Remote sensing maps record this information as an image. The classification of remote sensing maps is more complicated due to differences in sensors, imaging conditions, and types of information collected. They are also treated differently. The proposed method is suitable for visible light imaging and satisfies the remote sensing map of the vertical parallel projection imaging model. Considering the cost and simplicity of acquisition, this paper uses a satellite remote sensing map downloaded from the "satellite" mode of Google Maps. Its resolution accuracy is acceptable, it can be downloaded as long as it can be connected to the Internet, and it is completely free, so it is not only simple and practical but also low in cost. Its data components are more complex, mainly from Digital Globe and MDA Federal. Its imaging feature is a high-altitude bird's-eye view, which provides information on the roof of the building and covers a large area. In addition, it is visible light imaging and approximate vertical shooting; it can be assumed that the downloaded remote sensing image conforms to the vertical parallel projection imaging model; that is, it can meet the requirements of the algorithm.

3.2. Scene Visualization

3.2.1. Basic Steps of Reverse Tracking. According to the set image size, the number of rays is determined to be slightly higher than the total number of pixels in the image. If the image size is 160*120, the total number of pixels is 19200, and the light is 22630; if the image size is 640*480, the number of pixels is 307200, and the light is 363388; when the image size is 1024*1280, the number of pixels is 1310720, the light is 1550877, and the number of rays is more than 18% of the number of pixels, so that each cell has a light, and the extra light can be used to verify the correctness of each cell

calculation. When the number of rays is determined, tracking begins. The specific tracking process is as follows:

Step 1: Determine the position of the sensor and the viewing plane, and the light is directed to the scene through the viewing plane;

Step 2: When the light reaches the set opaque surface, the light source is tracked according to the surface reflection principle of the object. At this time, no other object is blocked between the light sources, and the reflection portion is a bright portion;

Step 3: When the light is reflected by the opaque surface, the scene entity that is set when tracking the light source is occluded (the sphere in the figure), and the reflective surface is dark;

Step 4: When the light hits the solid in the scene (the sphere in the figure), it is reflected to the opaque surface, and then the reflection can be traced to the light source. The reflection of the object is bright and there are some optics of the opaque surface characteristic; the light reverse tracking step is shown in Figure 1.

3.2.2. Acceleration Algorithm. In the process of ray tracing by POV ray, there are a large number of rays intersecting the object. In order to improve the judgment efficiency of the intersection of light and object, POV ray uses a variety of acceleration algorithms, including multiple layers nested bounding box algorithm. The most important multiple levels nested bounding box algorithm is introduced here, depending on the buffer algorithm and the ray buffer algorithm. The bounding box algorithm is widely used in ray tracing. The traditional bounding box algorithm divides the scene into virtual cubes (such as bounding boxes). First, it is determined which intersecting box the light intersects. If it intersects, it is judged whether the light and the entity in the bounding box are intersecting; compared with the original algorithm that uses image space as an order, this can greatly reduce the number of judgments and improve efficiency. The traditional bounding box algorithm steps are shown in Figure 2.

Since not every bounding box contains entities, POV ray uses a multiple layers nested bounding box algorithm. The multiple levels nested bounding box is similar to the structure of a tree, and the whole scene is first divided into larger bounding boxes. It is judged whether the light intersects with the bounding box. If they intersect, the intersecting bounding box is decomposed into smaller bounding boxes, and the judgment is performed again, and the layers are subdivided into multiple layers in order. In this way, when the entities are discretely distributed in the scene, the computational efficiency can be greatly improved; however, when the entities are continuously distributed, the efficiency is not as good as the traditional bounding box algorithm. Therefore, when using the multiple layers nested bounding box algorithm, the key is to start the bounding box. In POV ray, Bounding = on/off controls whether to use the bounding box, and Bounding Threshold = n controls the
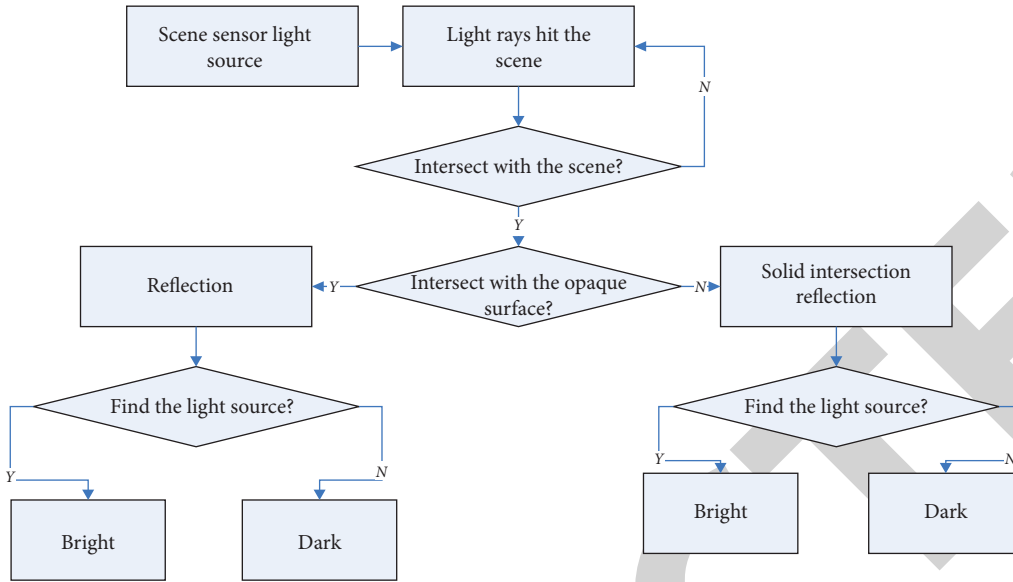
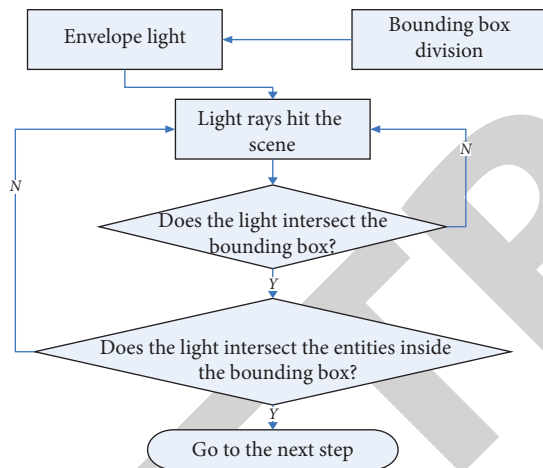Figure 1: Light reverse tracking step.

Figure 2: Traditional bounding box algorithm steps.

starting layer number n of the bounding box. POV ray defaults to 3 layers.

3.3. Generation of House Scenes.

The three-dimensional model of the house can be divided into two parts: the main body of the house and the roof. Therefore, the generation of the house scene can be divided into two parts: the main body of the house and the roof. Considering the complexity of the actual building, this article has been simplified accordingly: the complex building is broken down into a simple four-corner house model. For four-corner houses, there are roughly two categories, one is the most common flat-top houses, and the other is non-flat-top houses. To build a three-dimensional model of a house, you must first obtain some information about the house from some way. Undoubtedly, the information related to the house is mainly the corner coordinates of the house, the elevation of the bottom, and the height of the house; for nonflat houses, the information about the height of the roof of the house is also known.

3.3.1. Acquisition of Corner Information of Houses.

The house generally appears as a relatively regular shape in the two-dimensional image, so this paper decomposes it into a rectangle or a square, so that, for the house in the south of the south, it is only necessary to extract the coordinates of the two corners of the diagonal of the house, but for nonpositive south and north houses, the coordinates of the four corners of the house do not have a mutual relationship and must be extracted. In order to be able to generate the above two types of house models at the same time, this paper extracts all the corner coordinates of the house. So, how to extract corner coordinates has become a focus of this article. There are two ways to extract the corner information. One is to extract the boundary information of the house as a straight line segment and then find the intersection point of the straight line segment as the coordinates of the corner of the house; the other is to extract directly according to the gray information of the image (corner coordinates). By comparison, the latter method is found to be simpler and more accurate than the first method. When directly using the image gray scale to extract the angular coordinates of the house, a corner feature extraction algorithm such as a Harris operator and a Moravec operator is often used, which are all difficult to meet the needs of this study to extract the corner coordinates of the house. But with MATLAB's powerful matrix processing capabilities, you can easily solve the problem.

3.3.2. Acquisition of the Elevation of the Bottom of the House.

Using the classification map, we obtained the coordinate information of the four corner points on the bottom of the house. The acquisition of the elevation of the bottom of the

house is simpler. Since the bottom of the house in the study area is generally flat, the elevation of the bottom surface is uniform. The DEM elevation value of the selected area has been assumed to be 0 before, so the bottom elevation value is 0 from the DEM.

4. Discussion

4.1. Remote Sensing Image Registration Analysis

4.1.1. Registration Error.
The registration error is shown in Table 1 and Figure 3. It can be seen from the table that the camera optical axis position error is below 1.3 meters. The result of this accuracy is acceptable. The result has reached the practical standard and is large. In some cases, the resolution of the satellite remote sensing map is also meter level or lower.

The error of the registration algorithm mainly comes from two aspects, one is the error of the experiment itself, and the other is the algorithm error. For the omnidirectional diagram, errors in the relative positional mounting of the components of the camera, errors in the attitude of the device at the time of shooting (such as the level of the device), errors in the imaging device itself, etc. can cause errors in the imaging of the omnidirectional image. Remote sensing maps also have corresponding errors. The error of the experiment itself can be reduced by precisely adjusting the installation of the imaging device and correcting the obtained image. This part will not be discussed in this paper. The registration algorithm error mainly comes from feature extraction, because, from the principle of the algorithm, as long as the extracted features are absolutely accurate, the calculated registration result error is very small. Therefore, the accuracy of feature extraction is the bottleneck of the accuracy of the registration algorithm, especially the principal point registration method, because the features used are basically no information redundancy.

4.1.2. Registration Time Consuming.
The registration time consumption is shown in Table 2 and Figure 4. It can be seen from the table that both methods are relatively fast and can be controlled within 6 seconds. In comparison, the primary point registration method is faster. Moreover, the time-consuming search registration method will increase with the increase of scene buildings, and the main point registration method will be more time-consuming. This is consistent with theoretical analysis because the equal angle search registration method requires searching for a set of points. And the calculation amount used to judge the feasibility of any point in the point set increases with the increase of the number of linear features used and the principal point registration method only. One or several points need to be calculated, and the amount of calculation for calculating any point is fixed.

4.2. Scene Generation Results

4.2.1. Acquisition of Corners of Houses.
Therefore, based on the classification map obtained from the remote sensing image, the image is read into MATLAB using the imread function to generate a matrix, and then the matrix gray information of the house area is given to another matrix of the same size as the original image, so that these two matrices coincide. There are only two values: the gray value of the house and the background gray value. Each house is then marked or given a different color using the bwlabel function. Finally, the regionprops function can be used to extract the coordinates of the corner points of the region boundary. According to a certain rule, the coordinates of the four corner points of each house can be obtained. The corner points of the house extracted from the classification map are shown in Figure 5.

4.2.2. Generation of House Scenes.
When obtaining the building model construction information, the various types of information should be stored in the form of a matrix or an array, and then the patch function is used to draw the planes. The color of the roof and the color of the surrounding walls are separated, and the storage of the coordinate files is also required. The results of using MATLAB 3D building modeling are shown in Figure 6.

The modeling method of the three-dimensional house is relatively simple, and the required three-dimensional data is small, but the distortion is large, and there is a defect in the aesthetic angle when used for three-dimensional visualization. However, the modeling required in this paper can reflect the radiation characteristics of the ground object. Because the radiation characteristics are only related to the material of the ground under the conditions determined by the lighting conditions and the external environmental conditions, the radiation characteristics of the house are mainly the study of the radiation characteristics of both the exterior wall and the roof of the house. For this purpose, this modeling approach is feasible.

4.3. Clustering Algorithm on Remote Sensing Images.
Through experiments, it can be concluded that the clustering collective scale of ECUNGA algorithm is set to 5, 10, 20, 30, 40, and 50, the initial random parameter is 20, the maximum allowable algebra of GA is 500, and the maximum allowable stagnant algebra is 50. Perform clustering on the three data sets of Iris, Wine, and Glass, respectively, and compare the best correct rate, average correct rate of 20 times, and worst correct rate of the clustering results under the collective scale of each cluster. The experimental results are shown in Table 3 and shown in Figure 7.

From the experimental results, it can be found that the algorithm is not sensitive to the size of the clustering collective on the data sets Iris and Wine, while the clustering collective scale on the Glass data set is not very stable at 5 and

Table 1: Registration error.

	Isometric search registration method (m)	Main point registration method (m)
Test 1	1.2955	0.6411
Test 2	0.4324	0.6575

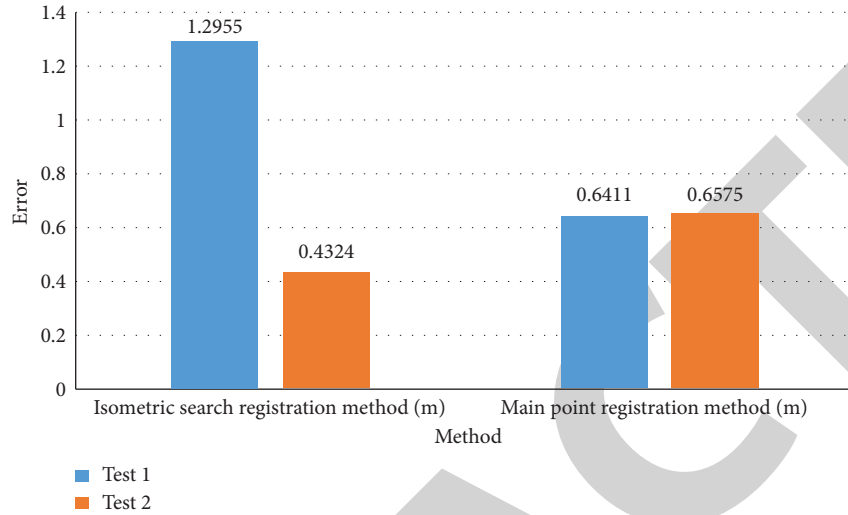Figure 3: Registration error.

Table 2: Registration time.

	Isometric search registration method (m)	Main point registration method (m)
Test 1	5.058	0.162
Test 2	1.813	0.114
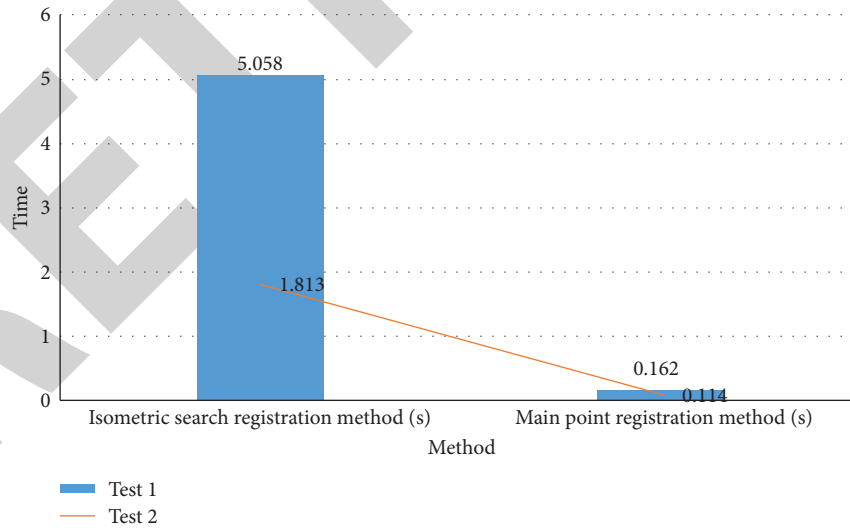
Figure 4: Registration time.

10, the best and worst results. The difference is relatively large.

4.4. Remote Sensing Image Registration. Analyze the experimental part and get Tables 4 and 5 and Figure 8.

Figure 9 shows the relationship between the number of sampling times of the RANSAC algorithm and the number of last correct matching point pairs and the execution time of the RANSAC algorithm. Red, green, and blue curves represent the experimental results on the test image pairs PA, PB, and PC, and the time unit is seconds. It can be
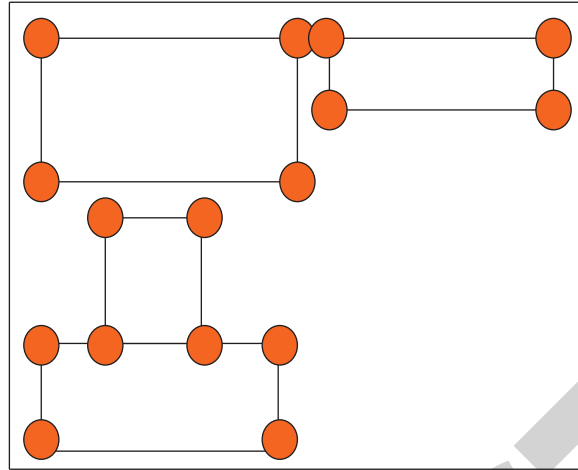
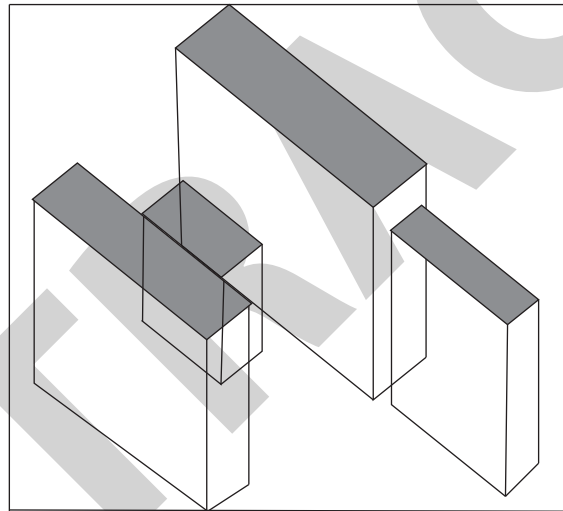Figure 5: House corner extraction effect map.

Figure 6: Housing scenario generation results.

Table 3: The influence of remote sensing image clustering algorithms.

Data set	Scale	Best	Average	Worst
Iris	All	0.89	0.89	0.89
Wine	All	0.7	0.7	0.7
	5	0.56	0.529	0.46
	10	0.54	0.48	0.46
Glass	20	0.54	0.52	0.52
	30	0.54	0.53	0.52
	40	0.54	0.51	0.52
	50	0.54	0.53	0.52

clearly seen from the figure that the number of correctly matched feature points remains unchanged after the number of iterative samples is greater than 80, so the threshold value is set to 100 in the RANSAC algorithm based on the similarity transformation model. RANSAC algorithm execution time and sampling times meet a linear relationship, so the RANSAC algorithm has a higher execution efficiency. It is worth noting that the RANSAC algorithm based on affine transformation usually requires more iterations. Since SIFT is a classic image registration algorithm based on scale space and point features, the SIFT algorithm is still compared here.

TABLE 4: The relationship between algorithm RANSAC sampling times and correct matching point pairs.

Sampling times	10	20	30	40	50	60	70	80	90	100
PA	2	8	8	8	8	8	8	8	8	8
PB	13	20	27	27	27	27	32	32	32	32
PC	76	76	76	76	76	91	91	91	91	91

TABLE 5: The relationship between algorithm RANSAC sampling times and running time.

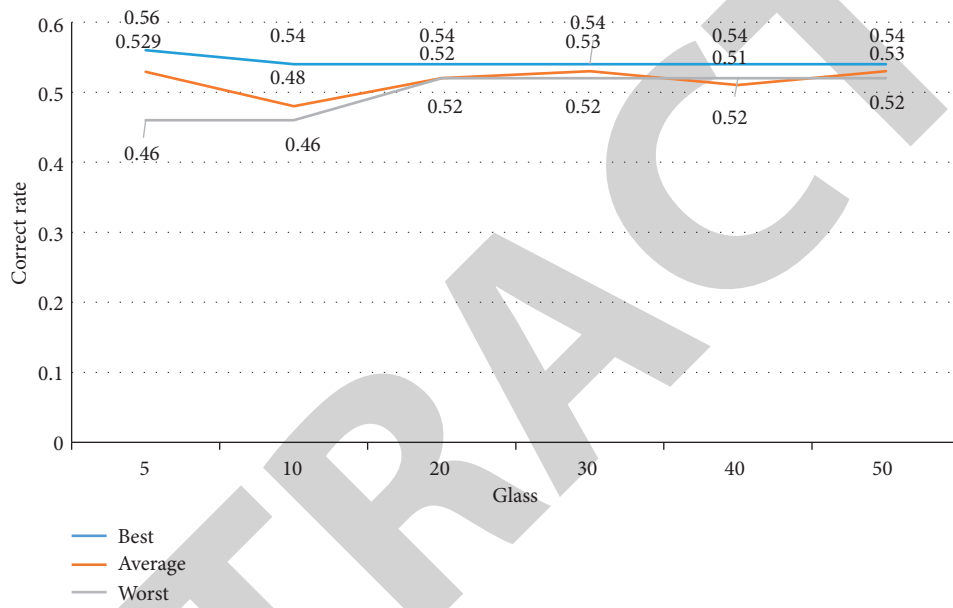| Sampling times	10	20	30	40	50	60	70	80	90	100
PA	0.05	0.021	0.026	0.039	0.04	0.045	0.047	0.06	0.07	0.08
PB	0.054	0.007	0.017	0.021	0.023	0.028	0.034	0.042	0.05	0.056
PC	0.045	0.017	0.021	0.031	0.034	0.037	0.041	0.054	0.063	0.08

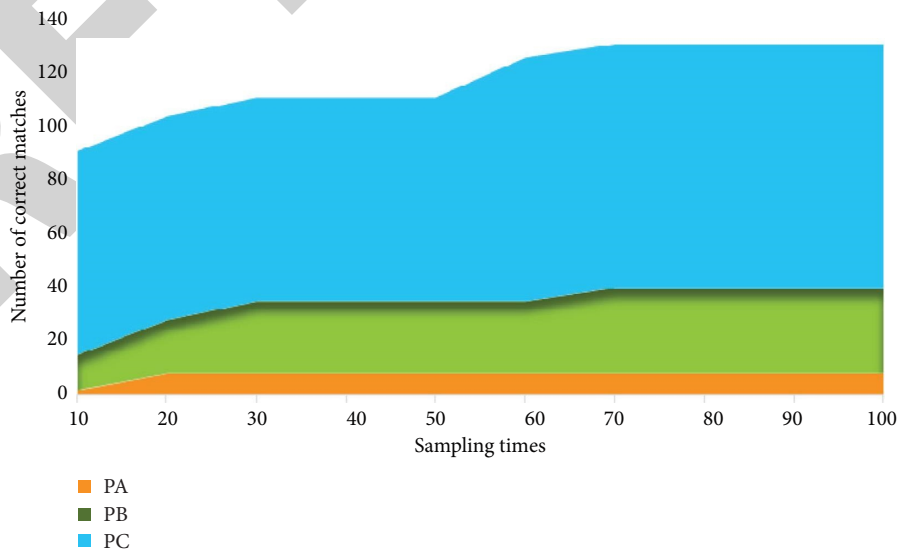FIGURE 7: The influence of remote sensing image clustering algorithm.

FIGURE 8: The relationship between the number of sampling times of the algorithm RANSAC and the correct matching point pair and running time.
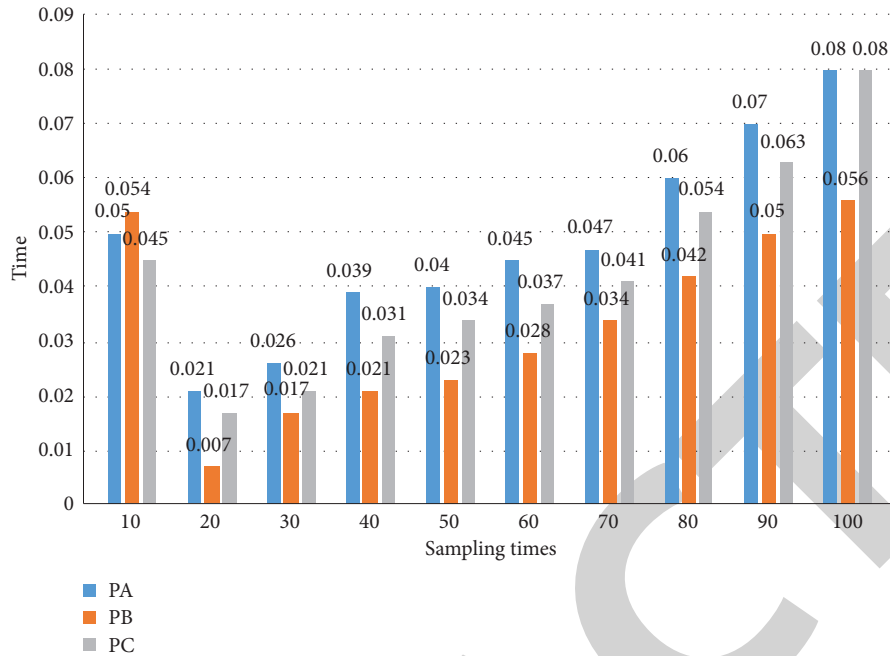
Figure 9: The relationship between the number of samples of the algorithm RANSAC and the running time.

It shows the image after registration and fusion using the algorithm proposed in this chapter. It can be seen that the edges and regions overlap well. Therefore, it can be judged intuitively that the registration result is accurate, which again proves the effectiveness of the algorithm proposed in this chapter.

5. Conclusions

This paper analyzes and elaborates the concept of hyperspectral remote sensing system and the absorption and reflection of electromagnetic waves. The structure of the hyperspectral scene system is analyzed, and the influencing factors of solar radiation and atmospheric effects and ground reflectivity model are introduced. The imaging mode and imaging principle of the imaging spectrometer were studied and discussed, and the parameters of the remote sensing system of the two scenes were determined.

For the simple scene, the simulation principle and implementation method of spatial correlation and spectral correlation in the ground reflectivity model are studied. The research shows that the spectral and spatial correlation of the features makes the spectral reflection of the pixels belong to different locations of the same type of features. The rate curve fluctuates around its mean reflectance. At the same time, the influence of the atmosphere on solar irradiation intensity and atmospheric transmission coefficient is analyzed. The analysis shows that the worse the visibility of the atmosphere, the smaller the solar irradiation intensity and the atmospheric transmission coefficient.

Based on the remote sensing image, a three-dimensional scene model is constructed. Based on the three-dimensional

scene model and the scene digital orthostatic image, a series of processing is extracted from the geometric and texture features of the scene orthostatic image and the contour features of the building. The minimum rule of the outline of the building is used to outsource the rectangle, and then these outsourcing rectangles are used as the constraint domain to segment the three-dimensional scene model to obtain the coarse single model. On this basis, the triangular and vertical classes of the triangular patches in the three-dimensional space are performed (classification). Calculate the roughness of the regional triangular patches. Based on this, combined with the height of the patch, the triangular patches that are not part of the building features in the coarse single model are purified, and then the adjacent patch growth method is used. The number of feature patches generates a more accurate building unit model and stores a single model from a single file, thus enabling automatic singular modeling of buildings in 3D scenes.

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

References

[1] A. Vanolo, "Smartmentality: the smart city as disciplinary strategy," *Urban Studies*, vol. 51, no. 5, pp. 883–898, 2014.

[2] P. Neirotti, A. De Marco, A. C. Cagliano, G. Mangano, and F. Scorrano, "Current trends in smart city initiatives: some stylised facts," *Cities*, vol. 38, no. 5, pp. 25–36, 2014.

[3] L. Sanchez, L. Muñoz, J. A. Galache et al., "SmartSantander: IoT experimentation over a smart city testbed," *Computer Networks*, vol. 61, no. 6, pp. 217–238, 2014.

[4] J. Jin, J. Gubbi, S. Marusic, and M. Palaniswami, "An information framework for creating a smart city through Internet of things," *IEEE Internet of Things Journal*, vol. 1, no. 2, pp. 112–121, 2014.

[5] Y. Li, W. Dai, Z. Ming, and M. Qiu, "Privacy protection for preventing data over-collection in smart city," *IEEE Transactions on Computers*, vol. 65, no. 5, pp. 1339–1350, 2016.

[6] F. Leccese, M. Cagnetti, and D. Trinca, "A smart city application: a fully controlled street lighting isle based on raspberry-pi card, a ZigBee sensor network and WiMAX," *Sensors*, vol. 14, no. 12, pp. 24408–24424, 2014.

[7] J. Gabrys, "Programming environments: environmentality and citizen sensing in the smart city," *Environment and Planning D: Society and Space*, vol. 32, no. 1, pp. 30–48, 2014.

[8] K. Czynska, "Application of lidar data and 3D-city models in visual impact simulations of tall buildings," *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 3, no. 7, pp. 1359–1366, 2015.

[9] R. Heno and L. Chandelier, "3D modeling of buildings: outstanding sites," *IEEE Transactions on Ultrasonics Ferroelectrics & Frequency Control*, vol. 50, no. 11, pp. 1429–1435, 2014.

[10] X. Zhang, Y. Lv, J. Tian, and Y. Pan, "An integrative approach for solar energy potential estimation through 3D modeling of buildings and trees," *Canadian Journal of Remote Sensing*, vol. 41, no. 2, pp. 126–134, 2015.

[11] H. Son and C. Kim, "Semantic as-built 3D modeling of structural elements of buildings based on local concavity and convexity," *Advanced Engineering Informatics*, vol. 34, no. 1, pp. 114–124, 2017.

[12] X. Zhang, L. Yang, and Y. Liu, "3D modeling of urban buildings and trees and its application in building-scale solar energy potential mapping," *Journal of Basic Science & Engineering*, vol. 22, no. 3, pp. 415–425, 2014.

[13] L. Díaz-Vilariño, K. Khoshelham, J. Martínez-Sánchez, and P. Arias, "3D modeling of building indoor spaces and closed doors from imagery and point clouds," *Sensors*, vol. 15, no. 2, pp. 3491–3512, 2015.

[14] M. Kedzierski and A. Fryskowska, "Terrestrial and aerial laser scanning data integration using wavelet analysis for the purpose of 3D building modeling," *Sensors*, vol. 14, no. 7, p. 12070, 2014.

[15] C. Huang and W. Bao, "A remote sensing image fusion algorithm based on the second generation curvelet transform and DS evidence theory," *Journal of the Indian Society of Remote Sensing*, vol. 42, no. 3, pp. 645–650, 2014.

[16] K. Sugihara and Z.-J. Shen, "Automatic generation of 3D building models by rectification of building polygons," *Advanced Science Letters*, vol. 21, no. 12, pp. 3649–3654, 2015.

[17] S. Rhee and T. Kim, "Dense 3d point cloud generation from uav images from image matching and global optimazation," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLI-B1, no. 1, pp. 1005–1009, 2016.

[18] S. G. Nebaba and A. A. Zakharova, "An algorithm for building deformable 3D human face models and justification of its applicability for recognition systems," *SPIIRAS Proceedings*, vol. 3, no. 52, pp. 157–179, 2017.

[19] J. Qin, C. Fang, Y. Wang, G. Li, and S. Wang, "Evaluation of three-dimensional urban expansion: a case study of Yangzhou City, Jiangsu Province, China," *Chinese Geographical Science*, vol. 25, no. 2, pp. 224–236, 2015.

[20] L. Ding, Y. Zhou, and B. Akinci, "Building Information Modeling (BIM) application framework: the process of expanding from 3D to computable nD," *Automation in Construction*, vol. 46, no. 6, pp. 82–93, 2014.