J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 2833--2856

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Research on Morphological Innovation in City Brand Image Based on Computer Graphics Technology

Siyang Liu[1, ✉]

[1] *Zhengzhou College of Finance and Economics, Zhengzhou, Henan, 450000, China*

ABSTRACT

Urban spatial structure and three-dimensional perspective can express personalized city brand image, which is an important feature of city brand form. In this paper, computer graphics technology is applied to design a city 3D modeling algorithm based on point cloud fusion, which transforms city information into city spatial visual symbols, and then carries out the innovation of city brand image morphology. Firstly, on the basis of binocular stereo vision, tilted image generation modeling technology is utilized to realize texture mapping 3D dense point cloud structure network. Aiming at the lack of accuracy of the sparse point cloud and the existence of noise points and mesh voids due to the influence of occlusion and shadows, we design the stereo vision PMVS algorithm based on the faceted slice in order to realize the densification of the point cloud. The algorithm performance is tested on the dataset using standard 3D reconstruction evaluation metrics F-score, chamfer distance (CD), and the application analysis of segmentation and merging execution efficiency for building clusters, optimization effect of rectangle fitting, and height calculation of building clusters, and the study finds that this paper's algorithm is ahead of the baseline model in 13 categories. When the number of regions reaches 70,000, the traditional RAG method takes 26.9 seconds, while this paper's algorithm only takes 14.8 seconds. The time consumption reduction reaches more than 40%. The average score of the aesthetic assessment of the city brand design is 83.47 points, and the 10 experts' evaluation of the spatial aesthetics is above 90 points, and the design is unanimously recognized. The study makes a useful exploration for the innovation of city brand image under the conditions of cutting-edge information technology.

Keywords: city brand image, 3D modeling, texture mapping, point cloud fusion

1. Introduction

In today's economic globalization, the global flow and allocation of capital and factors of production have made the competition for development resources more and more intense, and cities, as the core

✉ Corresponding author.
E-mail address: 18039102869@163.com (S. Liu).
Received 20 December 2024; Revised 05 February 2025; Accepted 25 March 2025; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-160

of countries and regions, occupy a very important position in the global competition [1-2]. Since the middle of the 18th century, developed European countries led by the United Kingdom have taken the lead in urbanization, and since then, more Western countries have gradually embarked on the road of urbanization, while China, as one of the largest developing countries, has only gradually entered the right track of urbanization in the early 50s of the last century [3-4]. Since the end of the 20th century, urbanization has been developing rapidly around the world, with developing countries growing significantly [5]. However, along with the development of urbanization, there is also the emergence of "urban problems". For example, the negative effects of "constructive destruction" and "urban beauty movement" such as vegetation destruction and environmental pollution. The phenomenon of "one side of a thousand cities" that has led to serious convergence due to the blind pursuit of high-rise buildings, wide roads, and big cities due to the "government performance should" has become an urgent problem to be solved in urban construction [6-8]. The image presented by the city will bring huge social and economic benefits to the city that attaches importance to image building, and create intangible assets that belong to the whole city [9]. Therefore, how to create a unique city brand image and enhance the original temperament of the city has become a hot topic in the academic circles at home and abroad, and more and more city managers have begun to realize the importance of building their own city brand image to enhance the city's personality.

Introducing the concept of “city brand image” in urban construction can not only make the city change from external development to internal development, enhance the civilization of the city, but also promote the economic development of the city, improve the quality of urban services and popularity, attract high-quality talents to settle down and tourists to travel [10-13]. Among them, city image and city brand are complementary and causal, shaping a good city image can create a better city brand, and the city brand is recognized and disseminated with the city image as a carrier [14-15]. Based on the theory of “city marketing” and “city image”, the proposed “city brand image” is a new concept of operating and managing the city, shaping the city image with the strategic method of city branding. The city image is shaped by the strategic method of city branding, thus forming the basic concept of city brand image [16]. The city is perceived existence, these perceptions mainly come from human six senses, they can make people get 80%-90% of the external information, and the visual occupies the vast majority, so in the development and construction of the city brand image, the construction of the visual image system is very important [17-18]. As a visual symbol, shape is an important part of the visual recognition system in the construction of city image, and its different presentations play a pivotal role in the construction of city brand image system [19]. The elements of city brand image include the city's characteristic products, people, scenery, architecture, historical heritage, culture, songs, dances, words, political economy and so on [20-21]. The resulting extension of the city brand image is diversified in the form of expression. However, exactly how to innovate the form is a problem to be solved in the construction of city brand image.

This paper first clarifies the value of 3D model narrative in city brand image construction, and then designs a city brand form innovation method based on 3D modeling algorithm. Aerial triangulation encryption is performed on the acquired UAV tilted image to determine the position and attitude of the UAV tilted image. Then the UAV tilted image with recovered aerial attitude and position is densely matched to generate the depth map, and finally the 3D model is generated and texture is mapped. Then for the problems of sparse point cloud model not realistic enough and generating mesh voids, the point cloud densification algorithm is constructed to solve the model optimization problem of the point cloud data, so as to make its 3D reconstruction model more fine and realistic. After applying the dataset to evaluate the performance of the algorithm, the algorithm is put into practical applications to check the quality of 3D modeling.

2. Image narratives and city brand image innovation

2.1. Concepts related to creative design of city brand image

2.1.1. Concept of city brand image. Every city needs a business card, in order to increase the visibility of the city, the core value of the city brand can bring more benefits for all members of this city. Establishing a good city brand image will have a rapid advancement of the city's future good modernization, improve the whole city's good cultural environment and the pride of the citizens, in the development of tourism, attracting talent, investment and other aspects play a vital role in the development of the city's competitiveness and accelerate the speed of economic development can be effectively improved. The establishment of the city brand is just like the brand of the product, a hundred years of famous brand standing is that it always adhere to the brand principles and brand unique value concept, the formation of differences and rivals, the city brand image is the same.

City brand is the value spirit that a city has the purpose to convey to the audience, the city's economy, culture, environment and other factors combined with the brand strategy constitutes the development strategy of the city brand. The establishment of city brand image is not only on the surface of “city marketing” and “city image”. The city brand image needs to have rich connotation, and the brand positioning and core value of the city should be integrated into the overall construction and development of the city.

2.1.2. Innovation in the form of city brand image:

1) City visual symbols. Doing a good job of city marketing, the real charm and value of the city can really be seen, and play a positive role in the economic development of the place. Creative design of city brand image is an important part of city marketing. Nowadays is the information age, people are more and more willing to receive information in a visual way. So the city brand image creative design should be based on the principle of visual communication design, the city's information into a visual language through color, graphics, symbols and so on to show [22].

2) City Identity. Nowadays, while the modernization of the city is fully emphasized, the importance of the city brand image creative design should not be ignored, and the city construction of a thousand cities makes the city culture gradually fade away, and the city brand image design is a way to show the cultural characteristics and values of the city. A complete set of city brand image design excellent city brand image design has allowed many cities to harvest valuable wealth value, so that the value of the city to experience the precipitation of time, will not be faded and lost because of the passage of time.

2.2. The Value of 3D Modeling Narrative in City Brand Image Building

2.2.1. Advantages of 3D modeling narratives:

1) Intuitive. With the development of science and technology culture, the form of narrative is getting richer and richer, and people's aesthetic demand has also changed, and three-dimensional model narrative has become the mainstream of narrative form because of its interesting and easy to understand. In this new media era, people's time in life is fragmented, do not have time to carefully read large sections of text to obtain information through understanding and analysis, the emergence of three-dimensional model narrative coincidentally make up for such a deficiency, the audience can intuitively understand the meaning of which want to express, and three-dimensional models bring the sense of the image and the sense of specificity can not be compared to the text of the narrative [23].

2) Co-temporality. Text narrative is told in accordance with the logical sequence of events, the story is expressed in a certain chronological order. Three-dimensional model narrative is in the narrative at

the same time can let people experience the time and space in the three-dimensional model, from the three-dimensional model of the audience's vision can feel in the space, from the narrative can understand the time clues of the occurrence of the matter, so through the three-dimensional model narrative people can clearly know the designer wants to express the spatial and temporal scope of the events that occur. Compared with the text narrative, 3D model narrative receives information more vividly by virtue of people's visual perception and imagination space. Shorten the psychological distance between the information and people, so that the audience is more immersive.

3) Aesthetic. Three-dimensional model narrative in addition to the narrative function, the composition of the work, color, proportion, and many other appearance factors constitute the beauty of the form from the appearance alone, with the aesthetic appeal of things. Compared with the narrative form of language and text, three-dimensional model narrative is often more attractive to people's eyes. The arrival of the new media era has intensified the three-dimensional model narrative in the appearance of the form of the development speed, people are increasingly using creative thinking in the design, so that the level of creative design to promote and improve, breaking the traditional form, three-dimensional model more visual impact.

2.2.2. Three-dimensional modeling narratives for city brand feature recognition. The use of three-dimensional model narrative to the creative design of city brand image is the focus of this paper, the term “three-dimensional model” has permeated the life and academic research, different from the image of the film, image narrative, three-dimensional model narrative tends to be more three-dimensional, a symbol represents a meaning and symbols, and in a certain cultural context will have a complete interpretation. A symbol represents a meaning and a symbol, which will be interpreted in a complete way in a certain cultural context. With three-dimensional visual symbols instead of textual symbols, the audience can perceive and experience the story and emotion that the designer wants to tell through such visual language. City brand image needs a systematic construction, in which a thematic cultural concept is needed as a support, so that the design of city brand image has a clear direction, and the theme in the narrative is the core content of the whole story. Just like the central idea of a movie, a piece of music or an article, the plot, rhythm or text description all revolves around the central point. The city's cultural characteristics, city values, city development status, etc. are covered in the theme, the theme in the narrative can make the city brand image design can coordinate the overall situation, 3D model narrative for the establishment of the city brand image play a role in grasping the overall direction.

3. City brand form innovation based on three-dimensional modeling algorithms

3.1. 3D Reconstruction of Drone Tilt Images

3.1.1. Main principles of binocular stereo vision. When the human eye acquires the 3D scene around it, the distance between the two eyes is B. When the two eyes focus on a point P on the surrounding features at a certain moment, the left eye and the right eye will each acquire an image containing that point. The position of point P can be approximated from the triangle formed by the position of the two eyes and point P. The following result is obtained by trigonometry:

$$\begin{cases} X_{left} = f\dfrac{x_c}{z_c} \\ X_{right} = f\dfrac{(x_c - B)}{z_c} \\ Y = f\dfrac{y_c}{z} \end{cases} \tag{1}$$

The above equation shows that the visual difference between the left and right photos is $Disparity = X_{left} - X_{right}$. The positional coordinates of point P in the camera coordinate system can be found by using the above visual difference Disparity between the left and right images:

$$\begin{cases} x_c = \dfrac{B \cdot X_{left}}{Disparity} \\ y_c = \dfrac{B \cdot Y}{Disparity} \\ z_c = \dfrac{B \cdot f}{Disparity} \end{cases} \tag{2}$$

This is obtained through the camera mapping relationship:

$$s_l \begin{bmatrix} X_l \\ Y_l \\ 1 \end{bmatrix} = \begin{bmatrix} f_l & 0 & 0 \\ 0 & f_l & 0 \\ 0 & 0 & l \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{3}$$

$$s_r \begin{bmatrix} X_r \\ Y_r \\ 1 \end{bmatrix} = \begin{bmatrix} f_r & 0 & 0 \\ 0 & f_r & 0 \\ 0 & 0 & l \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{4}$$

The left image space coordinate system $O - xyz$ is transformed by the spatial transformation matrix to obtain the right image space coordinate system $O_r - x_r y_r Z_r$, which is:

$$\begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = M_{lr} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} r_1 & r_2 & r_3 & t_x \\ r_4 & r_5 & r_6 & t_y \\ r_7 & r_8 & r_9 & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{5}$$

$$M_{lr} = [R \mid T] \tag{6}$$

In turn, the points of the left image space coordinate system can be converted to the right image space coordinate system according to the following equation:

$$\rho_r \begin{bmatrix} X_r \\ Y_r \\ 1 \end{bmatrix} = \begin{bmatrix} f_r r_1 & f_r r_2 & f_r r_3 & f_r t_x \\ f_r r_4 & f_r r_5 & f_r r_6 & f_r t_y \\ r_7 & r_8 & r_9 & t_z \end{bmatrix} \begin{bmatrix} zX_l / f_l \\ zY_l / f_l \\ z \\ 1 \end{bmatrix} \tag{7}$$

Thus, the three-dimensional coordinates of the space point can be expressed as:

$$\begin{cases} z = zX_l / f_l \\ y = zX_l / f_l \\ z = \dfrac{f_1\left(f_r t_x - X_r t_z\right)}{X_r\left(r_7 X_1 + r_8 Y_1 + f_1 r_9\right) - f_r\left(r_1 X_1 + r_2 Y_1 + f_1 r_3\right)} \\ \quad = \dfrac{f_1\left(f_r t_x - X_r t_z\right)}{X_r\left(r_7 X_1 + r_8 Y_1 + f_1 r_9\right) - f_r\left(r_4 X_1 + r_5 Y_1 + f_1 r_6\right)} \end{cases} \tag{8}$$

Therefore, as long as we know the left and right intra-photographic parameters/focal lengths f_r, f_i and the image coordinates of the spatial point in the left and right cameras, we are able to compute the 3D spatial coordinates of the measured point.

3.1.2. 3D texture mapping process. The dense point cloud can clearly show the outer contour of the target feature, but the details of the target feature are not well displayed. Therefore, it is necessary to 3D mesh the dense point cloud so that the point cloud becomes a 3D mesh with detailed structure, which in turn meets the needs of 3D texture mapping [24].

The model generated by the 3D mesh is basically readable, but it cannot clearly distinguish the distribution characteristics of each piece, especially where the shapes are similar. So 3D models need to be mapped with texture to improve the recognizability of 3D models. Texture mapping simply means projecting the 2D texture image into the 3D model to obtain a 3D visual model with sufficient realism, and then image processing the texture so that it can meet the requirements of people's visual interpretation.

3.1.3. Principles of airborne triangulation. The coordinates of the viewpoints are the observations x, y, and the azimuthal elements in the photograph are known, then the error of the beam method of null trilateration is expressed by the equation as follows:

$$\begin{cases} v_x = a_{11}\Delta X_s + a_{12}\Delta Y_s + a_{13}\Delta Z_s + a_{14}\Delta\varphi + a_{15}\Delta\omega \\ \quad +a_{16}\Delta\kappa - a_{11}\Delta X - a_{12}\Delta Y - a_{13}\Delta Z - l_x \\ v_y = a_{21}\Delta X_s + a_{22}\Delta Y_s + a_{23}\Delta Z_s + a_{24}\Delta\varphi + a_{25}\Delta\omega \\ \quad +a_{26}\Delta\kappa - a_{21}\Delta X - a_{22}\Delta Y - a_{23}\Delta Z - l_y \end{cases} \tag{9}$$

where $a_{11} = -\dfrac{f}{H}$, $a_{12} = 0$, $a_{13} = -\dfrac{x}{H}$, $a_{14} = -\left(f + \dfrac{x^2}{f}\right)$, $a_{15} = -\dfrac{xy}{f}$, $a_{16} = y$, $a_{21} = 0$, $a_{22} = -\dfrac{f}{H}$, $a_{23} = -\dfrac{y}{H}$, $a_{24} = -\dfrac{xy}{H}$, $a_{25} = -\left(f + \dfrac{y^2}{f}\right)$, $a_{26} = -x$, l_x, l_y are the constant terms of the equation in the x, y directions, ΔX, ΔY, ΔZ are the corrections to the coordinates of the point to be found, and ΔXS, ΔYS, ΔZS, $\Delta\varphi$, $\Delta\omega$, $\Delta\kappa$ are the elements of the exterior directions. Writing it in matrix form gives:

$$V = \begin{bmatrix} A & B \end{bmatrix} \begin{vmatrix} t \\ X \end{vmatrix} - L \tag{10}$$

Among them,

$$V = \begin{bmatrix} v_x & v_y \end{bmatrix}^T \tag{11}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \end{bmatrix} \tag{12}$$

$$B=\begin{bmatrix} -a_{11} & -a_{12} & -a_{13} \\ -a_{21} & -a_{22} & -a_{23} \end{bmatrix} \tag{13}$$

$$t=\begin{bmatrix} \Delta X_S & \Delta Y_S & \Delta Z_S & \Delta\varphi & \Delta\omega & \Delta\kappa \end{bmatrix} \tag{14}$$

$$X=\begin{bmatrix} \Delta X & \Delta Y & \Delta Z \end{bmatrix}^T \tag{15}$$

$$L=\begin{bmatrix} l_x & l_y \end{bmatrix}^T \tag{16}$$

The error equation corresponds to the normal equation:

$$\begin{bmatrix} A^T A & A^T B \\ B^T A & B^T B \end{bmatrix}\begin{bmatrix} t \\ X \end{bmatrix}=\begin{bmatrix} A^T L \\ B^T L \end{bmatrix} \tag{17}$$

Based on the above error equations and the corresponding normal equations the outer azimuth element t of each UAV tilt image can be derived, which in turn gives the position and attitude of each UAV tilt image.

3.1.4. Texture-mapped 3D dense point cloud structure mesh. For a graph $G=\langle V,E\rangle$ with sources s and sinks t, cut C is a subset of the set of edges E, utilizing cuts to partition the set of nodes V, into two parts. For a graph cut-construction network, the cut splits the triangular face sheet into two parts, surface and non-surface. The energy value of the energy equation corresponding to a cut is the sum of the energies of the edges interrupted by the cut, and the selection of the appropriate cut will result in the best 3D model surface. The energy function E represents a function of the energy change that occurs when a cut is selected, and the minimum value of the energy function is the optimal segmentation of the graph. Therefore, the energy function can be obtained using the optimal partition of the graph

The dense point cloud structure network problem can be converted into a graph cut problem. The graph cut problem needs to be solved by determining the energy equation corresponding to the solution of the graph. The relationship between neighboring nodes in the above graph corresponds to the smoothing term, and the relationship between general nodes and the origin and sink corresponds to the data term.

1) Energy term. The energy term of this paper focuses on describing the visibility of the tetrahedron of this node $E_{vis}(S)$. Each vertex in the original triangular mesh described above retains some visibility information, where multiple merges of vertices should contain visibility information for at least two constituent points. This information is decisive for the design of the tetrahedral surface visibility $E_{vis}(S)$ term; if a vertex belongs to the final surface, then it should be visible in the view from which it comes. Therefore, all tetrahedra where the light from the vertex intersects the center of the camera corresponding to this view should be marked as external, and tetrahedra behind this vertex should be marked as internal. However, there are some problems in this case, as the view changes, the number of rays increases, and for a tetrahedron there are several rays passing through it, and there is a contradiction that the tetrahedron is labeled as external by one ray and internal by another.

2) Smoothing term. The smoothing term consists of gloss consistency $E_{photo}(S)$ and surface smoothness $E_{area}(S)$. Gloss consistency $E_{photo}(S)$ measures how well a given surface S matches the different input images seen. In this paper, it is defined as the sum of the gloss consistency measurements ρ over the entire surface:

$$E_{photo}(S)=\int_S \rho dS=\sum_{T\in S}\rho(T)A(T) \tag{18}$$

The gloss consistency of each triangle can be computed by reconstructing it only in the view where its three vertices are reconstructed. For a triangle on a surface between the interior and exterior of the reconstructed object, its orientation needs to be taken into account. Gloss consistency means that the two possible orientations of a given triangle obtain different photo consistencies. This is easily mapped to the graph-cut framework, where for two adjacent tetrahedra, a weight $w_{Pq} = \rho_{\{\Pi i|\vec{d_i}\cdot\vec{n}>0\}}$ can be added to the triangular face T they share, where the normal vector $\vec{n}$ is the normal vector of the adjacent triangular face T (i.e., the line connecting the midpoints of even adjacent tetrahedra), and $\vec{d}$ is the orientation vector from the center of the triangular face to the center of the camera.

In addition to the photo consistency there is a smoothing term that also includes the surface smoothness energy:

$$E_{area}(S) = A(S) = \int_S dS = \sum_{T\in S}(T) \tag{19}$$

For each pair of neighboring tetrahedra, there is a common triangular face between them. At two neighboring points in the figure, a weight of $w = A(T)$ is added to the edge connecting these two points in this paper, and similarly, we give the same weight to the direction opposite to the edge that has been this one. In this way, this paper encourages the minimization of the area of the triangular faces that make up the surface.

Based on the above energy terms, it is possible to write the energy equation for the graph cut problem corresponding to this graph:

$$E(S) = E_{vis}(S) + \lambda_{area}E_{area}(S) + \lambda_{photo}E_{photo}(S) \tag{20}$$

where this equation consists of three terms, tetrahedral surface visibility $E_{vis}(S)$ is the data term of the energy equation, and surface smoothness of the triangular faces between neighboring tetrahedra $E_{area}(S)$ is the smoothing term of the energy equation. λ_{area} is the two positive weights.

Through the energy equation, this paper finds the minimum cut of the graph corresponding to the energy equation by finding the minimum energy. In this paper, light visibility, photo consistency, and surface smoothness are used as the energy term and smoothing term of the equation to determine the degree of consistency between the model and the real features, and then the energy equation is used to simulate the process of the static features recognized by the human eye. Then the optimal segmentation of the initial 3D mesh is solved by the minimum cut of the graph mentioned above, and the noise points and noise surfaces in the triangular mesh are removed by the optimal segmentation of the initial 3D mesh, thus realizing the accurate 3D conformal mesh of the dense point cloud.

3.2. Model optimization for point cloud data

3.2.1. Characterization of multi-view point cloud based data. A point cloud is discrete data that describes the structural and morphological properties of the surface of a target object, and each point in the point cloud data is regarded as the most basic geometrical entity, which can be used to describe the target object of any structure. By processing the point cloud data of the target object, the structural information of its surface can be extracted to realize the 3D point cloud reconstruction of the target object [25]. The point cloud data acquired by different devices will be presented in different forms, which are generally categorized into the following four arrangements:

1) Scattered point cloud: the point cloud data presents a disordered and scattered distribution, and no topological relationship can be seen. Generally through the camera shooting, drones, laser scanning point cloud will appear in this form.

2) Scanning line point cloud: the point cloud presents a semi-scattered state, and the topological relationship of some points can be seen. Generally, line structure light scanning is used to obtain such point cloud data.

3) Polygonal point cloud: all the point cloud data are connected with each other in the form of a grid. Some magnetic resonance imaging equipment can obtain polygonal point clouds.

4) Display point cloud: the point cloud data are arranged in orderly rows and columns like a matrix, so that the topological relationship of the points can be clearly seen.

The point cloud generated based on multi-view is generally a scattered point cloud, in which the point cloud data contains 3D coordinates and RGB information, and its point cloud data has the following characteristics:

1) The multi-view based point cloud is acquired by shooting images with a camera, and its equipment is simple and convenient to acquire, less costly and more efficient, and easy to learn.

2) The point cloud based on multi-view is scattered, and the distribution of this point cloud helps to generate the surface texture of the target object.

3) The point cloud generated based on multi-view contains or can calculate rich point cloud information, such as 3D spatial coordinates, RGB, texture information, etc., which provides a strong support for the 3D reconstruction of the point cloud later.

4) Although the accuracy of point cloud acquisition based on multi-view has been getting higher and higher, the generated point cloud is still sparse compared with radar data, and in the process of acquiring images, due to the influence of some problems such as occlusion, shadows, and the environment, the point cloud data may have noisy points, grid voids and other phenomena.

3.2.2. Point cloud data preprocessing. In the process of obtaining the point cloud through close-range photogrammetry, the influence of the surrounding environment and instrument error when obtaining the point cloud will inevitably cause some outlying noise points, which will have an impact on the accuracy of the model and the modeling speed. These outliers are the points that are at a certain distance from the main point cloud data, and the number is less than the main point cloud. The main point cloud usually contains a lot of main information, while the outlier noise points contain less point information. Therefore, we need to ignore the outlier information that is far away from the main point cloud, and these outlier points will have some impact on the accuracy and modeling speed of the model.

3.2.3. Dense point cloud reconstruction. The sparse point cloud does not have more detailed information about the buildings, and the dense matching of the point cloud is needed to get the densified point cloud. The PMVS algorithm, which is based on faceted stereo vision, has a relatively good densification effect and a wide range of applications, so the PMVS algorithm is used to densify the sparse point cloud for reconstruction.

1) PMVS thickening algorithm. PMVS algorithm is a multi-view stereo visualization densification reconstruction based on facets. Firstly, SIFT feature points are detected on the graph, and the initial facets are generated through feature point matching, and then the dense spatial point cloud is obtained through several iterations of facets, filtering, and expansion.The flow of PMVS algorithm is shown in Fig. 1.

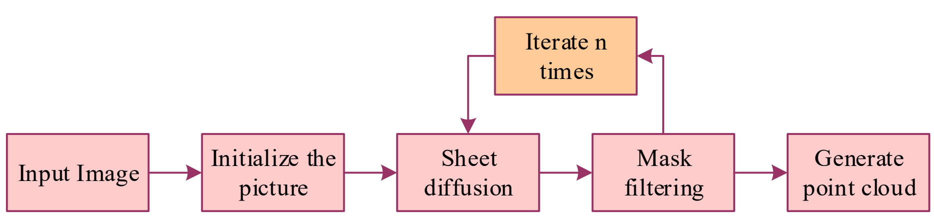

Fig. 1. PMVS algorithm process diagram

(1) Facet modeling. A facet is actually a tangent plane to a local surface. A slice is defined as p, $c(p)$ is the center of slice p, $n(p)$ represents the unit normal vector perpendicular to slice p and pointing to the optical center of the camera, and the corresponding reference image is denoted by $R(p)$.

(2) Gray scale consistency. Firstly, the set of images of facet p is represented by $V(P)$, and each facet is divided into a grid of $\mu\times\mu$. The angle threshold is set to $\tau\left(\tau=\frac{\pi}{3}\right)$. After bilinear interpolation sampling method, the pixel gray values $q\left(p,I_1\right)$ obtained by the projection of facet p on the image I_1 are captured, and then the normalized inter-correlation data of $q\left(p,I_1\right)$ and $q\left(p,I_2\right)$ is subtracted by 1. The gray consistency function of images I_1 and I_2 is denoted as $h(p,I,R(p))$.

The gray level consistency function $g(p)$ is shown in the following equation:

$$g(p)=\frac{1}{|V(p)/R(p)|}\sum_{I\in V(p)/R(p)}h(p,I,R(P)) \tag{21}$$

This algorithm has relatively high requirements on the image, and in the actual shooting process often can not achieve the ideal situation, will be obscured by some obstacles, so the function $g(p)$ will no longer be applicable. In order to solve this problem, it is necessary to limit the threshold of the image gray consistency function a. $V*(P)$ is used to represent the ensemble of images in $V(p)$ that meet the gray consistency, expressed as follows:

$$V*(P)=\{I\mid I\in V(p),h(p,I,R(p))\le a\} \tag{22}$$

Replace $V(p)$ with $V*(P)$, after which the following equation is obtained:

$$g(p)=\frac{1}{|V*(p)/R(p)|}\sum_{I\in V*(p)/R(p)}h(p,I,R(P)) \tag{23}$$

P of the actual set of visible images $V*(P)$ can be derived by the above equation (23).

The PMVS (patch-based multi-view stereo) algorithm is to reconstruct at least one facet in each image block, and the reconstruction process of facet is divided into the following three steps:

(1) Initialization Matching. Initialization matching is done by feature point detection and matching, so as to obtain a series of sparse facets. Generally, the matching of PMVS thickening algorithm uses the facets generated by the motion recovery structure (SFM).

(2) Expansion of facets. Facet expansion as the name suggests is the process of repeatedly generating new facets around an existing facet. For a facet p, which corresponds to an image block $C(x,y)$, a point cloud facet is obtained after the feature point detection as described above, and since the facet is too sparse to fully cover the image block, it is necessary to extend the facet so that each image block $C(x,y)$ contains at least one projection of the facet. The expansion steps are as follows:

a) Determine whether the image block needs to be extended or not. Given a facet p, the set of image blocks of neighboring images in the image is denoted by $C(p)$. Set $C_i(x,y)$ to be a certain image block and $C_i(x,y)\in C(p)$. If a facet p' is already included in $C_i(x,y)$ and facet p' is adjacent to the defined facet p, there is no longer any need to extend the facet to it, and $C_i(x,y)$ is directly removed from $C(p)$. The adjacency of a facet is shown below:

$$\left|\left(C(p)-C\left(p'\right)\right)\cdot n(p)\right|+\left|\left(C(p)-C\left(p'\right)\right)\cdot n\left(p'\right)\right|<2\rho \tag{24}$$

b) Expansion of facets. For an image block $C_i(x,y)$ in a neighboring set of image blocks $C(p)$, an extension is used to generate a new facet p'.

(3) Filtering of facets. In the process of face sheet expansion, some wrong faces are often generated in the process of generating a large number of faces, so it is necessary to filter some faces with weak gray scale consistency and geometric consistency. If the conditions of the following equation cannot be met, then we need to remove the face sheet p. Because when the face sheet p is outside the surface, $V*(p)$ and $1-g*(p)$ are very small, so we can filter out the face sheet.

$$\left|V^*(p)\right|(1-g*(p)) < \sum_{p_i \in U(p)} 1-g*(p_i) \tag{25}$$

2) PMVS densification algorithm implementation. Using the above facet-based stereo visual densification (PMVS) algorithm to complete the dense matching computation through the three stages of facet matching, expansion and filtering, to obtain a dense 3D point cloud of the target object, and finally expand the sparse dense point cloud into a dense 3D point cloud data. Obviously, the number of point clouds after densification increases exponentially, which better reflects the details of the model, and then supports the innovation of city brand image design form under the 3D perspective.

3.3. City brand image form innovation based on 3D modeling

3.3.1. Characterization of urban 3D scene representation. Cognition refers to the analogy of the human brain to a computerized information processing system and includes the processes of information acquisition, processing and subjective information generation.

The process of spatial cognition covers five stages, and Figure 2 shows the process of each stage. Firstly, the conceptualization of the real world forms the conceptual space, after digitization and abstraction, the data space is obtained, the information space rules obtained from further processing and organization and analysis are processed to constitute the knowledge space, and finally put into application practice to form the cognitive space.

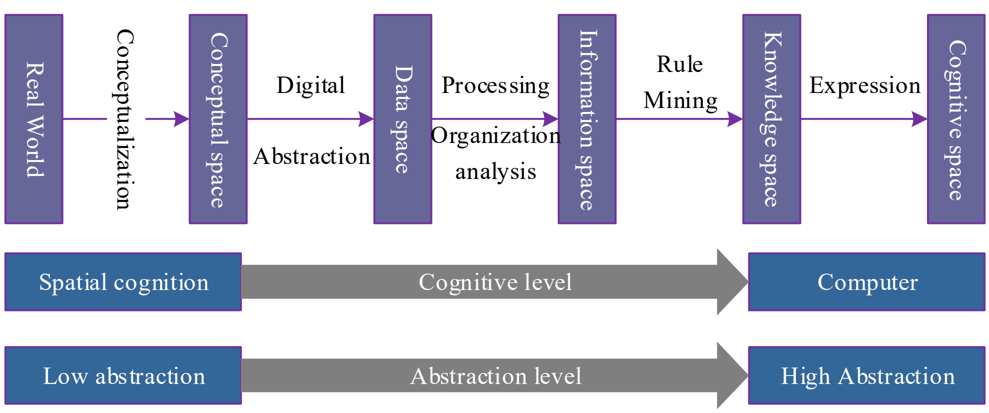

Fig. 2. Spatial cognition

In the urban 3D scene, interest and attention refers to people's interest or preference for a model, and importance refers to the significance of a model's role in the 3D scene, such as whether it is a landmark building, the supporting role of the urban distribution framework, and the prominence of people's visual law, the interest and attention to a high level of the region to a certain extent reflects the importance of the region is high, and the importance of a high level of the 3D A building model or an area with high importance can often get more attention and interest. Interest, attention, building importance in people's actual production has practical guidance significance, so that people based on cognitive needs to actively recognize and explore things, manifested as a strong selective and positive response to the thing, stirring people's attention, leaving the impression of the memory is more profound.

As a more intuitive form of spatial information expression, the urban 3D scene transmits different levels of spatial information to the audience through 3D data such as building models in a way that conforms to the laws of people's visual cognition, delaying the loading of modeling data that people don't pay attention to or hiding them in a place where visibility is not required, while retaining and highlighting important features for better understanding and communication. In the browsing process, people tend to have a progressive cognitive process from whole to local and from rough to detailed, combining the audience interest, attention and importance of 3D model in the 3D scene, in the overall view, the first perception is the overall spatial distribution characteristics, road network location correlation and density characteristics of the city 3D model, and in the further browsing, the volume characteristics, structural characteristics and positional importance of the city 3D model When further browsing, the volume characteristics, structural characteristics and location importance of the city 3D model become salient factors, and when observing individual geographic elements, its salient key points are more likely to attract people's attention and concern.

3.3.2. Multidimensional Mapping of City Brand Image by 3D Modeling. The word "mapping" is often used in computer science and mathematics. In the concept of mapping, the corresponding object is called "image", and the corresponding object is called "original image", which means that the state of the "original image" is reflected through the "image". On the basis of the above explanation, in the relationship between the 3D model and the city brand image, the 3D model of the city can be understood as the "image", and the city brand image can be understood as the "original image", which is a set of mapping relationships. As a three-dimensional model in urban communication, the urban space entity itself is not only a physical entity, but also a symbol carrier that generates urban meaning and enriches the urban image, and has a mediating role in the process of urban communication, especially in the brand city, which can reflect the multi-dimensional urban brand image. It should be pointed out that the mapping of the 3D model to the city brand image is not a one-to-one correspondence, but the relationship between the 3D model and the interweaving and fusion of the city brand image, and the relationship between the 3D model mapping the city brand image is shown in Figure 3.

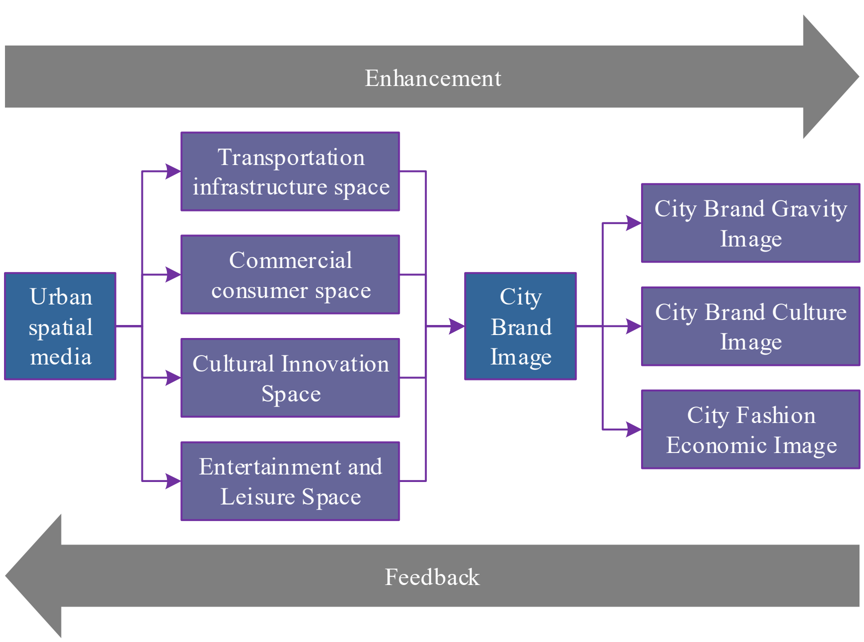

Fig. 3. The relationship diagram of the city brand image

1) Mapping the city's brand gravity image. The vitality and sustainability of urban development comes from the attraction of the city to talents, resources and other elements of urban development, and one of the important paths to build a brand city is to shape the gravitational image of the city brand. As a city entity, 3D model is not only a basic condition for city development, but also can map the city brand gravity image through 3D model. The city brand gravity image is expressed in all aspects of the city brand image, but the most important is the attraction of the transportation base and the attraction of the city brand concept.

Throughout the development of the world's brand capital, the city's transportation infrastructure space is an inextricable topic. The brand city has a very modern comprehensive transportation system, such as subway, bus and other three-dimensional urban public transportation system, external modern high-speed railroads, international airports and other developed transportation network trunk, so that the brand city and the region, between the international establishment of accessible liaison transportation system, improve the brand city of a strong trade, culture and other radiation, so that the city and the continued to maintain the brand The city will continue to maintain the vitality of the brand city. For example, Paris, the capital of the city, has a developed metro system and urban and suburban and neighboring cities of the regional rapid transport trunk lines, making Paris, an international metropolis in the developed transportation network trunk lines, Paris is also known as the “Queen of the City” more rich in brand charm.

The city brand concept is also an important part of the city's image of attractiveness. City brand concept can be said to be the brand city's spiritual temperament, is based on the city's comprehensive aspects of the city's unique style and habits. As a medium with communication function, urban space embodies the urban community's identification with the brand culture symbols and the attitude system of the urban community in the brand space of the city, which becomes an important symbol carrier for the outside world to perceive and understand the unique brand image of the city. For example, the Bund and the Lujiazui urban space shaped the scene of Shanghai as a modern internationalized metropolis, and conveyed the city's brand concept of “pluralistic, integrated and open metropolis”.

2) Mapping city brand culture image. City is a container for culture gathering, preservation and inheritance, and it is the source and driving force for a city to have brand temperament, shaping the brand culture image and creating the unique brand culture influence of the city are the important paths for the development of the city to branding.

Brand city is a new form of city that we study, it is a city presented with cultural brand, and it is the ideal of urban development and the goal to chase. The spatial entity in the city contains the unique cultural heritage of the city, which is an important cultural representation of the brand city, and is the static cultural symbol and language of the brand city. On the one hand the sample of urban space, the textual meaning and image of urban space are characterized according to the cultural heritage of the city, which not only condenses the deep cultural heritage of the city, but also coincides with the evolution of tolerance and innovation in the process of the city's development towards branding and modernization, and all the macroscopic and microcosmic changes and iterations in the society such as the social relationship and lifestyle will be presented within the urban space. Throughout the world famous brand cities - Paris, London, Tokyo, Rome, Hong Kong, etc. are the birthplace of international brand culture, and their urban space and city brand culture are closely related mapping relationship. For example, the Louvre in France, with its classic architectural space style and mellow artistic connotation, presents to the world a French romantic and innovative brand culture atmosphere, which makes this architectural space one of the cultural landmarks of the brand city of Paris. For brand cities, urban space is a symbolic space that unites the city's unique temperament and culture.

3) Mapping the economic image of the city brand. The material basis of human social practice is the state of economic conditions, similarly, the development and evolution of the city is also closely related to the economic state. From the point of view of international brand city characteristics, the world's brand capitals are basically distributed in developed countries, Paris, London and other five brand capitals have become the trend center of the world's brands, and the ultimate determinant of its status is economic strength.

For brand cities to gain strong urban competitiveness and economic development vitality, they must have strong economic development space and economic consumption space. Here the city space as a medium, is a kind of analytical carrier to deal with the relationship between two things, the city's brand economic development space and the city's consumption space here is more than an ordinary intermediary, but as a three-dimensional model to become the city brand economic image can be shown and readable text.

3.3.3. Innovative direction of city brand image based on 3D modeling:

1) Design city logo. City logo can vividly and graphically reflect a city's cultural characteristics and humanistic style, usually combined with the city's iconic buildings, influential figures and urban culture to design a city logo, integrate the city features, and design an anthropomorphic animation image in line with the city's positioning, so as to enhance people's impression and cognition of the city. The city logo based on the 3D city model can play a significant role in promoting the city. Like city logo, city souvenir is also an effective publicity medium for city brand image. Identifying and selecting the most representative city landmarks as 3D models to be made into city souvenirs can achieve the effect of city publicity. It is also possible to design matching gift boxes of souvenirs that fit the city's characteristics and have a beautiful appearance. At present, the city souvenir appearance packaging design more successful case is the Beijing “Forbidden City” architectural complex model souvenirs, “Forbidden City” Museum souvenirs most of the high-precision model of the Forbidden City, the Drum Tower and other famous historical and cultural architectural complexes in Beijing, accompanied by engraved and printed Most of the souvenirs of the “Forbidden City” Museum are packaged with high-precision models of the Forbidden City, Drum Tower and other famous historical and cultural buildings in Beijing, accompanied by exquisite souvenir packages engraved with the “Forbidden City” logo.

2) Designing city souvenirs. Like city symbols, city souvenirs are also an effective publicity medium for city brand image. Identifying and selecting the most representative city landmarks as 3D models to be made into city souvenirs can achieve the effect of city publicity. It is also possible to design matching gift boxes of souvenirs that fit the city's characteristics and have a beautiful appearance. At present, the city souvenir appearance packaging design is more successful case of Beijing “Forbidden City” architectural complex model souvenirs, “Forbidden City” Museum souvenirs most of the high-precision model of the Forbidden City, the Drum Tower and other famous historical and cultural architectural complexes in Beijing, accompanied by engraved and printed The “Forbidden City” word mark of the exquisite souvenir packaging. This design not only makes the souvenirs highly ornamental, but also publicizes the city of Beijing in an intangible way during the circulation of the souvenirs, so that its urban cultural characteristics are gradually known to the public. Therefore, the city souvenirs symbolize and inherit the city culture and help the public understand the city's humanistic flavor.

4. Algorithm performance and 3D modeling quality

4.1. Assessment of urban modeling data sets

Implemented using python language, all the experiments are carried out in Ubuntu16.04, i9-10350H, GTX4090 environment. The effectiveness of the algorithm in this paper is verified by comparing it with the current mainstream dimensional reconstruction algorithms.

4.1.1. Data sets and assessment criteria. A 3D reconstruction dataset containing 50k 3D CAD models in 13 categories was used. Each model was rendered from 24 randomly selected viewpoints, and the internal and external parameters of the rendering camera were recorded. The models in this dataset are partially derived from the ShapeNet dataset.

Standard 3D reconstruction evaluation metrics were used for the experiments. Vertices were first uniformly sampled from the reconstruction results and the surface of the real mesh model to form a predicted model point cloud and a real model point cloud. Then the Precision and Recall are calculated by calculating whether the vertices in the predicted point cloud can match the vertices in the nearest real point cloud within a certain threshold range to determine whether each vertex is predicted correctly. Then, the F-score is calculated as a combined evaluation of the accuracy and recall, which is essentially a harmonic average of the accuracy and recall.

In addition, the chamfer distance (CD) between the model and the real is also used as an evaluation metric. For each vertex in a point set, the chamfer distance algorithm finds the closest neighboring vertex in another point set and sums the squares of the vertex distances. If the chamfer distance is considered as a function of the spatial position with respect to the two point sets, the chamfer distance is continuous and segment-smoothed. Since the chamfer distance is computed independently for each vertex, it can be computed in parallel. Strictly speaking, CD is not a distance in Euclidean space, since the triangle inequality does not hold for the chamfer distance. Even so, the chamfer distance still gives a good response to the gap between all vertices in the two model spaces, and the smaller the value the more accurate the 3D reconstruction.

4.1.2. Comparison of experimental results. This paper compares the current mainstream deep learning based 3D reconstruction algorithms: 3DR2N2, LSM and Pix2Mesh. since 3DR2N2 and LSM produce voxel models, they need to be meshed using the moving cube algorithm, and then vertices are uniformly sampled from the surface of the model to obtain a point cloud model for evaluation. For quantitative comparison, 3D reconstruction is performed using this paper's algorithm, 3DR2N2, LSM and Pixel2Mesh respectively with the same dataset and evaluation metrics are calculated.

Table 1 shows the F-score comparison results of the results of each algorithm. The matching threshold is set to 1. The F-score value of this paper leads the rest of the baseline algorithms in all 13 categories, with an average value of 62.11, which is higher than the 47.27, 49.94, and 60.15 of 3DR2N2, LSM, and P2M.

Table 1. Comparison of the results of the match threshold setting 1

	3DR2N2	LSM	P2M	This model
Couch	45.75	43.52	52.26	53.16
Cabinet	54.65	51.47	60.74	61.19
Bench	45.4	49.57	57.81	59.29
Chair	38.16	48.79	54.55	55.47
Monitor	36.9	43.86	51.85	52.8
Firearm	56.49	57.19	74.33	77.14
Speaker	41.69	45.69	49.71	54.03
Lamp	32.52	46.32	48.16	53.31
Cellphone	59.24	61.31	70.71	71.22
Plane	48.69	56.06	72.52	72.92
Table	49.58	49.18	66.93	69.75
Car	60.94	52.18	68.59	69.55
Watercraft	41.26	48.59	56.19	57.63
Mean	47.27	49.94	60.15	62.11

Table 2 shows the F-score comparison of the reconstruction results of each algorithm when the matching threshold is set to 2. The average F-score of the baseline model 3DR2N2, LSM, P2M, and this paper's algorithm are 63.17, 65.95, 75.3, and 79.22, respectively. This paper's algorithm is still ahead of the baseline model in 13 categories.

Table 2. Comparison of the results of the match threshold setting 2

	3DR2N2	LSM	P2M	This model
Couch	60.76	56.09	71.2	74.35
Cabinet	64.75	61.02	78	82.04
Bench	62.94	66.66	73.17	78.72
Chair	54.96	65.56	71	74.62
Monitor	49.47	56.94	68.21	75.53
Firearm	77.65	74.38	85.02	87.54
Speaker	53.18	56.94	66.09	69.7
Lamp	50.24	65.67	61.87	72.56
Cellphone	70.92	72.04	83.63	85
Plane	71.69	77.55	82.98	85.88
Table	63.66	62.75	80.76	84.25
Car	79.19	69.4	85.19	88.46
Watercraft	64.48	68.09	71.04	71.22
Mean	63.17	65.95	75.3	79.22

In summary, it is concluded that this paper's algorithm outperforms other algorithms under the vast majority of categories of objects, and the advantage of this paper's algorithm is even more pronounced after relaxing the evaluation criteria (the matching threshold becomes twice as high). It is worth noting that the 3DR2N2 algorithm does not use the parameters inside and outside the camera during operation, while the Pixel2Mesh algorithm can only perform monocular image 3D reconstruction.

Further, Table 3 shows the results of the chamfer distance (CD) comparison between the reconstructed model and the real model for each algorithm. The average CD of this paper's algorithm is 0.518, which is lower than the rest of the algorithms, demonstrating a 3D modeling performance that is closer to the real model.

Table 3. Comparison of CD values for each algorithm

	3DR2N2	LSM	P2M	This model
Couch	0.641	0.721	0.456	0.4
Cabinet	0.29	0.64	0.339	0.219
Bench	1.274	0.538	0.62	0.501
Chair	0.741	0.493	0.578	0.416
Monitor	0.917	0.558	0.749	0.429
Firearm	0.271	0.391	0.42	0.229
Speaker	1.306	0.778	0.741	0.506
Lamp	6.896	1.784	1.287	1.137
Cellphone	0.703	0.327	0.397	0.282
Plane	0.382	0.492	0.465	0.282
Table	1.095	1.008	0.48	0.363
Car	0.312	0.329	0.273	0.239

Watercraft	0.474	0.489	0.653	0.445
Mean	0.955	0.641	0.596	0.518

4.2. Practical 3D modeling applications

4.2.1. Efficiency of cluster-oriented split-merge implementation. In this paper, the labeled images of segmented high-resolution remote sensing images of an urban area with 7649 regional units are used as experimental objects. The method described in this paper and the traditional RAG and NNG methods are used for comparison experiments to compare the efficiency of region merging execution.

Table 4 records the execution efficiency of the three methods at different scales. Where the initial length of the hierarchical queue is the initial number of neighboring chains of the hierarchical queue under the corresponding scale. As can be seen from the table, under the same scale, the RAG algorithm has the longest queue length, NNG is the second longest, and this paper's method is the shortest. This is because this paper's algorithm follows the three necessary conditions for hierarchical queue access, which makes the queue length much shorter, and its length is only 2%~15% of RAG and 30%~60% of NNG. Secondly, at the smallest scale, all the three methods need to perform the construction of relevant data structures and numerical initialization operations before merging, but the merging thereafter can directly utilize the results obtained from the previous scale without reconstruction and initialization, so the speed of the later merging will be increased. From the total time consumption, it can be obtained that the execution efficiency of the NNG method is 3.64 times higher than that of the RAG method, while the execution efficiency of the method in this paper is 11.7 times higher than that of the RAG method.

The above experiment proves that the algorithm of this paper can minimize the length of rank queue and significantly improve the merging speed under the guarantee of segmentation accuracy from the perspective of segmentation and merging efficiency.

Table 4. Regional merge execution efficiency comparison

Scale	RAG		NNG		This model	
	Class queue initial length	Time (s)	Class queue initial length	Time (s)	Class queue initial length	Time (s)
500	20708	8.483	4072	2.818	2016	1.959
100	17499	4.273	1243	1.761	559	0.104
1500	15852	4.086	672	1.134	287	0.092
2000	13842	3.942	324	0.867	185	0.065
2500	10764	3.185	274	0.549	105	0.049
3000	8973	3.007	113	0.276	54	0.036
		26.976		7.405		2.305

From the previous section, it can be seen that the construction, retrieval and sorting of the hierarchical queue is the key link in the region merging method, which directly affects the efficiency of region merging. A city high-resolution remote sensing image containing blue, green and red bands with a spatial resolution of 0.5 m is selected as an experimental object, and an over-segmented labeled image is obtained by using the Watershed Transform before merging, and then the segmentation accuracy is compared with the merging efficiency.

Figure 4 shows the relationship between the number of regions and the merging time consumption, and it can be found that with the increase of the number of regions in the initial segmentation of the image, the advantage of the merging speed of the method in this paper also increases gradually. Especially when the number of regions reaches 70,000, the time consumption of the RAG method is

26.9 seconds, while the algorithm in this paper is only 14.8 seconds. The reduction of time consumption reaches more than 40%, which is undoubtedly a great advantage for processing large amount of remote sensing image data.

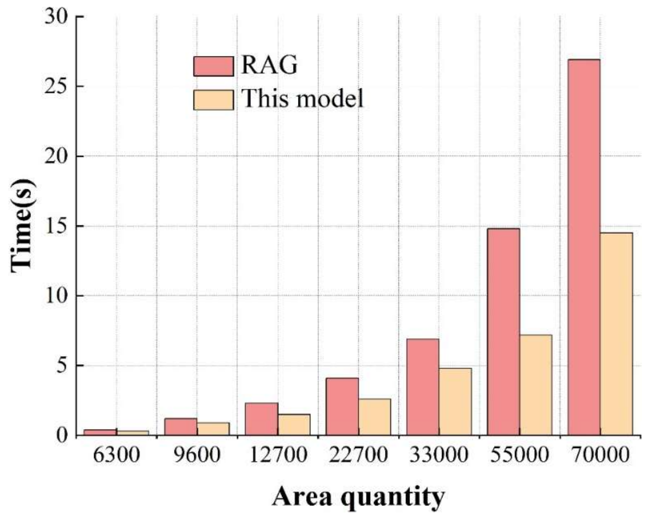

Fig. 4. The efficiency comparison of the two methods

Figure 5 shows the relationship between the cost spacing and the relative segmentation accuracy (RAS).The RAS is obtained by the following method: assuming that the segmentation junction under the traditional method is the correct result, and comparing the result obtained by this paper's method, the percentage of correctly segmented pixels is the RAS.As seen in the figure, with the reduction of the cost spacing, the RAS is gradually improved, and when the cost spacing is less than 10, the RAS can be maintained at a high level of over 93%. 93% or higher, the accuracy of this paper's algorithm can meet the needs of city brand image innovation based on 3D model narrative.

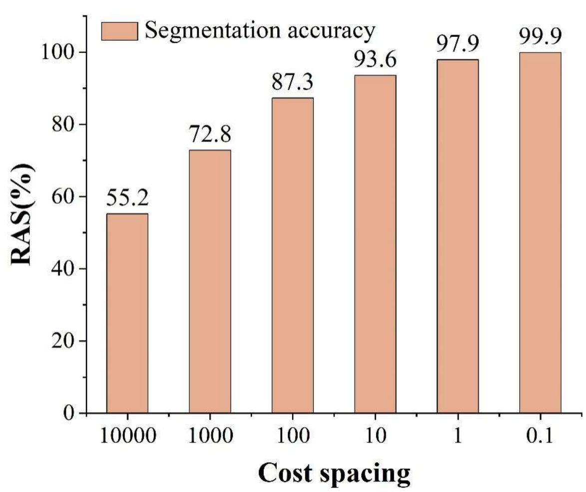

Fig. 5. The relationship between cost spacing and relative segmentation accuracy

4.2.2. Rectangular fit optimization for building clusters. Taking the initial vectorized graph of a group of buildings as an example, the algorithm of this paper is used to conduct experiments, and the area errors of the original polygon and the final fitted polygon are counted, and the statistical results are shown in Table 5. 8 polygons, the percentage of the area error is 1.61% at the largest, 0.44% at the smallest, and 0.88% at the average.

Table 5. Area error statistics

N	Original polygon area	Pseudopolyvariant area	Area error	Percentage of error
1	461.81	463.82	2.01	0.44%
2	557.19	560.18	2.99	0.54%
3	386.198	389.39	3.192	0.83%
4	532.94	528.45	4.49	0.84%
5	456.46	463.82	7.36	1.61%
6	517.34	511.39	5.95	1.15%
7	453.88	459.22	5.34	1.18%
8	619.29	616.35	2.94	0.47%
				0.88%

The algorithm in this paper achieves a more satisfactory result both in terms of morphology and area accuracy. The results have reached a high level of accuracy for large urban building groups, which can meet the application requirements of city brand image innovation design based on 3D model narrative.

4.2.3. Calculation of the height of the complex. Taking the façade information of a city landmark building group as the object of detection, the heights of all the buildings in the group are calculated. Ten buildings were randomly selected for height error detection, and the results are shown in Table 6. The average error is 0.8 m, and the maximum absolute value of error is 1.37 m. It can be seen that the method of this paper can achieve better height extraction accuracy.

Table 6. Statistical table of high quality error sampling

N	Model calculation height	Actual height	Error absolute
1	49.06	47.94	1.12
2	51.98	53.34	1.36
3	18.43	18.62	0.19
4	39.96	41.07	1.11
5	21.85	23.62	1.77
6	22.86	22.79	0.07
7	37.65	36.28	1.37
8	18.28	18.66	0.38
9	53.39	54.71	1.32
10	20.16	20.38	0.22
			0.80

5. City brand image innovation - the case of City Z

Applying the three-dimensional modeling method and city brand image innovation path proposed in this paper, city brand form innovation is carried out for Z city. Firstly, three-dimensional modeling of drone tilted image is carried out for City Z. Then, three-dimensional city brand image elements are refined and selected for the design of city logo, new media pictorial, and architectural complex model souvenirs. Finally, city brand image experts were invited to evaluate the design.

5.1. Innovative design of city brand image

5.1.1. Refinement of 3D model design elements. After completing the 3D modeling of City Z, the design elements are refined. The figurative buildings selected based on the criteria of humanistic significance, appearance characteristics, popularity and transportation convenience cannot be directly refined as city brand image elements, because the style and size ratio of each building are different, which will lose the sense of unity and aesthetics. If the design elements of each feature prominence, color clutter, the proportion of random differences without a neat sense of harmony, then the overall combination of the picture will appear chaotic, lose harmony, difficult to become a unified whole. If the elements are all neat and uniform but no difference, the whole picture will appear monotonous, hard, rigid, and lack of dynamic sense of art.

Similar colors are used to express the sense of unity and wholeness of the architectural complex, and the differences between the complexes are expressed through enhanced features. Since the methods of refining are similar, Nanshan and Zhenjiang Municipal Library are used as examples to discuss in detail the refining process of diverse and unified figurative elements. Exaggeration is the exaggeration of the characteristics of the figurative architectural elements, Nanshan's stone pagoda has a very obvious appearance, divided into three parts, each of which has an upward flying eaves, the highest in the middle, the two sides correspondingly lowered, presenting a “convex” form, with four columns interspersed in it. After grasping these characteristics of Nanshan stone memorial hall, we can draw the concise shape of Nanshan memorial hall on the basis of three-dimensional modeling.

Ultimately, this paper based on three-dimensional modeling to select and refine, can represent the history, culture and humanism of Z City, the symbol of the architectural complex or natural landscape: Maoshan Pai Fang, Dashikou building, museum building, Baotashan Pagoda.

5.1.2. Morphing of design elements. The deformation of the elements should conform to certain rules, and on the premise of not violating the harmony of the overall surface painting, carry out continuous and rhythmic changes based on the three-dimensional model. The arrangement and organization of basic shapes should be carried out under the guidance of certain rules, including the direction and position of the arrangement of basic shapes. In the innovative design of the city brand image of City Z, the picture is enriched by the evolution and deformation of the architectural and humanistic elements, and the changes in the color and shape of the elements are integrated in each stage of evolution, which finally form a harmonious and unified picture.

5.1.3. Reconfiguration of design elements. Refactoring is the process of going from decentralization to centralization. It is the design process that is diametrically opposed to dispersal. It means to build, to reorganize, to shape. Reconstruction is a new merger and reorganization of the physical properties of the elements. Specifically manifested in the elements of the color, shape of the collocation again, so that the design elements to form a complementary relationship between each other, and as a contrast.

For example, the representative building of Da Shi Kou is a modern high-rise building, which represents the modernization achievements of Z city. In order to show the towering character, its design should not only consider the height of the building itself, but also consider the contrast with other buildings. There is a huge difference between ancient buildings and modern buildings in terms of height, and this huge difference will destroy the harmony in the design. Therefore, in the design of the large-scale buildings such as Dashikou, you can add some details according to its own characteristics.

After the refinement, deformation and reconstruction of the design elements of the three-dimensional model, we successfully designed the city logo, new media pictorial and architectural model souvenirs containing the humanistic and spiritual elements of the city of Z, such as the Maoshan pagoda, the Dashikou building, the museum building, and the pagoda on Baotashan Mountain.

5.2. City brand image innovation design effect

Ten city brand image experts were invited to assess the quality of the innovative design of the city brand image. The assessment was divided into 3 dimensions: cultural recognition, design aesthetics, and narrative elements of the 3D model, and each expert scored each item out of 100 points.

Table 7 shows the results of the assessment of cultural recognition. The cultural recognition contains 4 assessment items: symbolism, traditional characteristics, spiritual connotation and story connotation, and the average score given by 10 experts is 82.55, which is at a high level. The average scores of symbolic meaning, traditional characteristics, spiritual connotation and story connotation were 79.60, 82.61, 82.70 and 85.30 respectively, and the cultural recognition of the innovative design was recognized by the experts.

Table 7. Evaluation of cultural recognition

N	Symbolism	Traditional art	Spiritual connotation	Story connotation	Mean
1	80.55	84.53	88.41	88.18	85.42
2	78.83	87.81	88.62	84.68	84.99
3	76.98	85.56	83.14	88.86	83.64
4	70.96	77.65	86.41	89.89	81.23
5	81.82	74.78	87.13	82.83	81.64
6	77.76	84.64	74.96	89.83	81.80
7	86.12	75.83	76.27	87.80	81.51
8	82.72	88.23	78.89	85.65	83.87
9	83.04	80.96	81.22	78.40	80.91
10	77.18	86.06	81.99	76.85	80.52
Mean	79.60	82.61	82.70	85.30	82.55

Table 8 shows the results of the design aesthetics assessment. The average scores of stylistic expression, color pattern, spatial aesthetics, and structural innovation are 83.04, 77.66, 93.11, and 80.07 respectively, with an overall average of 83.47. Thanks to the accurate reproduction of the humanistic architectural complexes such as Maoshan Paikou, Dashikou Building, Museum Building, and Pagoda on Baotashan Mountain by the 3D modeling algorithm designed in this paper, the evaluation of the spatial aesthetics of the 10 experts is above 90 points, and the design is unanimously recognized.

Table 8. Design aesthetic evaluation results

N	Style expression	Color pattern	Spatial aesthetic feeling	Structural innovation	Mean
1	74.03	84.93	90.03	81.12	82.53
2	85.76	74.76	96.16	71.8	82.12
3	78.59	76.67	94.75	81.44	82.86
4	82.37	84.19	91.9	71.8	82.57
5	88.38	76.81	92.68	89.42	86.82
6	82.02	72.06	92.94	80.06	81.77
7	78.21	71.27	93.47	80.99	80.99
8	87.35	76.04	90.54	84	84.48
9	86.58	79.81	92.89	82.98	85.57
10	87.08	80.09	95.73	77.09	85.00
Mean	83.04	77.66	93.11	80.07	83.47

Table 9 shows the evaluation results of narrative elements. In terms of landmarks, the narrative scores for the three-dimensional model of Maoshan Paikou, Dashikou Building, Museum Ocean Building, and Baotashan Pagoda were 78.04, 79.58, 83.08, and 78.17. In terms of emotional intention, the scores of cultural confidence, historical heritage, and sunny life were 81.29, 80.70, and 81.41 respectively. The overall average score for narrative elements was 80.32.

Table 9. Narrative element assessment results

N	Landmark building				Affective intention			Mean
	MAO shan	Grand city	Museum floor	Pagoda mountain	Cultural confidence	Historical background	Sunshin e life	
1	71.35	76.79	83.63	70.52	75.41	82.86	83.65	77.74
2	71.82	78.29	79.53	79.02	80.12	77.43	82.89	78.44
3	71.12	83.81	83.33	84.79	89.21	87.7	79.78	82.82
4	71.18	85.76	74.3	82.52	83.14	85.56	76.1	79.79
5	84.15	79.24	86.72	70.27	80.25	70.57	73.55	77.82
6	87.37	82.25	89.27	71.02	87.93	81.48	88.21	83.93
7	76.33	73.31	76.23	80.42	82.57	84.42	86.13	79.92
8	81.72	81.28	89.34	87.6	72.28	71.18	87.88	81.61
9	75.7	82.37	80.41	74.53	86.76	85.01	85.4	81.45
10	89.65	72.7	88.08	81.05	75.19	80.74	70.54	79.71
Mean	78.04	79.58	83.08	78.17	81.29	80.70	81.41	80.32

For the city brand image innovation design of City Z in this paper, the average scores of experts' assessment in the three dimensions of cultural recognition, design aesthetics, and narrative elements of 3D modeling are all above 80 points. It can be considered that the city brand design innovation based on 3D modeling algorithm proposed in this paper is effective.

6. Conclusion

In this paper, a city 3D modeling algorithm based on point cloud fusion is constructed for the realization of 3D city visual symbols to enhance city recognizability and innovate city brand image. The 3D modeling performance of the designed algorithm is verified on the public dataset, and the city brand image design innovation is carried out with City Z as an example, and the following conclusions are drawn:

1) When the matching threshold is set to 1, the F-score value of this paper is ahead of the rest of the baseline algorithms in all 13 categories, with an average value of 61.27, which is higher than that of 3DR2N2, LSM, and P2M, which are 47.27, 49.94, and 60.15.

2) The average chamfer distance of the algorithm in this paper is 0.518, which is lower than the rest of the algorithms, showing a 3D modeling performance closer to the real model.

3) In the segmentation and merging execution efficiency test, the execution efficiency of the NNG method is 3.64 times higher than that of the RAG method, while the execution efficiency of this paper's method is 11.7 times higher than that of the RAG method. The algorithm in this paper can minimize the length of hierarchical queue and significantly improve the merging speed under the guarantee of segmentation accuracy.

4) In summary, the algorithm in this paper can perform city 3D modeling more accurately and efficiently, can meet the needs of city brand image innovation based on 3D model narrative, and the quality of innovative design is recognized by experts, and can be used in the practice of city brand image innovation.

References

[1] Rossi, U. (2017). Cities in global capitalism. John Wiley & Sons.

[2] Kirk, R. (2024). Neoliberal necropolitics and the global competition for urban dominance. Geoforum, 155, 104107.

[3] Henderson, J. V., & Turner, M. A. (2020). Urbanization in the developing world: too early or too slow?. Journal of Economic Perspectives, 34(3), 150-173.

[4] Wei, H., Wei, & Li. (2019). Urbanization in China. Springer Singapore.

[5] Haase, D., Güneralp, B., Dahiya, B., Bai, X., & Elmqvist, T. (2018). Global urbanization. The Urban Planet: Knowledge Towards Sustainable Cities, 19, 326-339.

[6] Hou, Y., Li, J., Li, G., & Qi, W. (2023). Negative effects of urbanization on plants: a global meta-analysis. Ecology and Evolution, 13(4), e9894.

[7] Liang, L., Wang, Z., & Li, J. (2019). The effect of urbanization on environmental pollution in rapidly developing urban agglomerations. Journal of cleaner production, 237, 117649.

[8] Yan, H., & Liu, Z. (2023). A new perspective on the evaluation of urbanization sustainability: urban health examination. Sustainability, 15(12), 9338.

[9] Beinart, J. (2021). Image construction in premodern cities. In Imaging the City (pp. 3-32). Routledge.

[10] Lee, C., Ra, H., Oh, Y., & Lee, C. (2021). Global Busan City brand image development strategy-SWOT/AHP analysis. East Asian Journal of Business Economics (EAJBE), 9(3), 115-124.

[11] Kim, D. H. (2018). The effects of city brand image on city brand recognition and city loyalty. The Journal of Industrial Distribution & Business, 9(3), 69-79.

[12] Park, G., & Park, S. (2018). A study on perceived city brand image and consumer behavior. Journal of Digital Convergence, 16(7), 145-151.

[13] Cleave, E., & Arku, G. (2020). Immigrant attraction through place branding? Evidence of city-level effectiveness from Canada’s London. Cities, 97, 102502.

[14] De Luca, G., & Shirvani, A. (2019). Boosting city image for creation of a certain city brand. Geographica Pannonica, 23, 23-31.

[15] Pompe, A. (2019). DESIGNING THE IMAGE AND THE PERCEPTION OF THE CITY AND ITS'BRAND: THE IMPORTANCE AND IMPACT OF QUALITATIVE URBANISTIC ELEMENTS. Advances in Business-Related Scientific Research Journal, 10(2).

[16] Ma, W., de Jong, M., Hoppe, T., & de Bruijne, M. (2021). From city promotion via city marketing to city branding: Examining urban strategies in 23 Chinese cities. Cities, 116, 103269.

[17] Thampanichwat, C., Meksrisawat, P., Jinjantarawong, N., Sinnugool, S., Phaibulputhipong, P., Chunhajinda, P., & Bhutdhakomut, B. (2024). A Systematic Review of Architecture Stimulating Attention through the Six Senses of Humans. Sustainability, 16(15), 6371.

[18] Li, Y. (2023, July). Brand Visual Identification System for Urban Brand Image Design. In International Conference on Frontier Computing (pp. 46-52). Singapore: Springer Nature Singapore.

[19] MATLOVIČOVÁ, K., TIRPÁKOVÁ, E., & MOCÁK, P. (2019). City brand image: Semiotic perspective a case study of Prague. Folia Geographica, 61(1), 120.

[20] Özbey, D. Ö., & Başer, H. H. (2015). A Study about The Role of The Elements of City Brand Image in City Branding Process: The Example of Konya. The Macrotheme Review-A multidisciplinary journal of global macro trends, 4(7), 179-190.

[21] Chan, A., Suryadipura, D., & Novel, N. J. A. (2022). Brand element: exploring the effect on city branding. International Journal of Professional Business Review, 7(4), e0750-e0750.

[22] Xuchang Zhang & Wenqiang Wu. (2024). The Influence of Folk Sports Tourism on the Construction of City Brand Image. Applied Mathematics and Nonlinear Sciences(1),

[23] Qingxiang Li,Lingege Long,Xinwei Li,Guidong Yang,Chenhang Bian,Benyun Zhao... & Ben M. Chen. (2025). Life cycle cost analysis of circular photovoltaic façade in dense urban environment using 3D modeling. Renewable Energy121914-121914.

[24] Adriano Dos Reis Carvalho,Matheus Viana da Silva & Cesar H Comin.(2024).Artificial vascular image generation using blood vessel texture maps..Computers in biology and medicine109226.

[25] ZeyuZhang,DongLiang,HaibinHuang & LuSun. (2024). Virtual trial assembly of large steel members with bolted connections based on multiscale point cloud fusion. Computer-Aided Civil and Infrastructure Engineering(17),2619-2641.