Journal of Computers Vol. 31 No. 4, 2020, pp. 14-27
doi:10.3966/199115992020083104002

A Virtual-real Fusion Method of an Augmented Reality Map Based on a Hierarchical Semantic Marker

Zhangang Wang*, Xian Zhu

School of Information and Communication Engineering, Beijing Information Science and Technology University, Beijing, China
wangzg@bistu.edu.cn, zhuxian@bistu.edu.cn

Received 26 November 2018; Revised 10 August 2019; Accepted 11 October 2019

Abstract. With the continuous development of augmented reality technology research and the deepening of map and geographic information applications, an increasing number of scholars are trying to apply mobile augmented reality technology to map display enhancement and geographic information visualization. At present, research of augmented reality maps faces challenges. The first challenge is lack of a specific marker for map space elements; furthermore, a marker without a semantic description and location characteristics is only used for augmented reality registration. In addition, there is a lack of three-dimensional (3D) tracking registration and virtual reality fusion for planar maps, which is key to achieving spatial consistency and enhanced expression of augmented reality maps. To solve the above problems, a new user-defined hierarchical semantic marker in the map space that combines AR mark and QR code information is designed and proposed. The user-defined hierarchical semantic marker in the map space changes the fixed markup mode in the traditional augmented reality. It can represent the map elements' spatial locations, semantic information, and multi-dimensional attribute information with flexible features and universality. At the same time, the marker can be used to assist the virtual reality fusion and model matching, and lay a foundation for the application of the augmented reality map. This paper introduces a city map user-defined hierarchical recognition method, explains its principles in detail, and expounds on the process of 3D registration of real scenes and virtual model in space. Applying the registration method based on computer vision, four types of coordinate systems, i.e., the real space coordinate system, the virtual space coordinate system, the viewpoint coordinate system, and the image plane coordinate system, are unified to realize seamless fusion of the real scene and virtual information; then, a virtual reality fusion algorithm for map models is proposed. The virtual models corresponding to the user-defined markers are seamlessly integrated with the map in reality. Finally, results of conducted experiments and analyses of map marker hierarchy recognition, map virtual-real information fusion, etc., reveal that the virtual-real fusion between the user-defined map marker and map algorithm in this study is effective. The marker contains a wealth of information, the virtual models are displayed accurately, and the effect of enhanced expression is achieved. An analysis of the efficiency and recognition rate of the user-defined map marker hierarchy recognition reveals that the map marker hierarchy recognition method proposed in this paper, with a high recognition accuracy and a degree of universality, is fast, and its recognition efficiency meets the real-time requirements of augmented reality.

Keywords: 3D registration, augmented reality, hierarchy recognition, map semantic marker, virtual-real fusion

1 Introduction

Augmented reality maps (ARMs) combine the advantages of flat maps with a strong sense of substance and easy reading, and are multi-dimensional, dynamic, interactive, and realistic. They break through the

* Corresponding Author

expression of static visualization of traditional flat maps, provide a variety of human-computer interaction methods, and enhance the interactive perception of spatial information and multidimensional dynamic visualization of the real world [1-2]. The application of augmented reality technology in mobile terminals is called mobile augmented reality (MAR). Mobile augmented reality maps promote integration of traditional flat maps and mobile terminals to produce better application and expression effects. With the rapid development of mobile augmented reality technology and the deepening of the application of maps and geographic information, mobile augmented reality has become a research hotspot [3-4]. An increasing number of scholars are trying to apply augmented reality technology to map display enhancement and geographic information visualization [2, 5-6].

The core challenges in mobile augmented reality map research are map target tracking, registration, and model virtual-real fusion. Augmented reality requires registering a real scene with virtual information in a three-dimensional (3D) space to determine the mapping relationship of virtual objects in different coordinate spaces. The matching process is called registration [7]. According to different tracking technologies, 3D tracking and registration can be achieved with two methods: computer vision-based tracking registration and sensor-based tracking registration. Among them, the computer vision-based tracking registration method is the most commonly used, and it can be divided into marked tracking and unmarked tracking.

In marked tracking technology, artificial images or makers as identification marks are usually placed on a map object or boundaries of interest in a real-world environment [2]. Users capture images of the paper map through cameras in mobile devices and track positions of map surfaces, and then use the pattern recognition technology to identify preset single marks. In this way, users' positions or viewpoints are determined, and finally 3D DEM, buildings, virtual models, such as of rivers, or information are superimposed onto paper maps [6]. Virtual information and paper maps are integrated together to form an enhanced map [8], and users can interact with the maps to query or measure certain map attributes in their own way [1]. Currently, the marked tracking method is widely applied, but the function of markers is relatively simple. Usually markers are only used for augmented reality registration of specific objects [2], and they are recognized and parsed for a single marker. They lack universality and are without semantic description and positional characteristics. At present, to the best of the authors' knowledge, there are no general marker studies on map space elements.

After realizing 3D tracking registration of map targets, virtual models or information corresponding to map targets needs to be directly superimposed onto the real-world scene to realize seamless integration of the virtual world and reality. Research on virtual-real fusion of the mobile augmented reality map mainly involves consistency processing and showing the rationality of virtual models and real world objects. Its purpose is to make the virtual model and information appear correctly and naturally on the camera screen so that the model and the surrounding environment can correspond with each other. In the process of virtual-real fusion of a mobile augmented reality map, it is necessary to comprehensively consider the occlusion relationship between the virtual model and the map in the real environment and the position, angle, and size of the model display [9]. Through visualization technology, relevant geographic information can be transformed into virtual models, such as DEM, annotation, map symbols, and video [10]. Using the result of map target tracking and registration, the virtual model can be accurately located and merged in the scene of a real map [11] to achieve the enhanced expression effect of virtual-real fusion. The mobility and content diversity of mobile augmented reality maps increase the difficulty of virtual reality. The virtual model in a traditional virtual-real fusion algorithm needs to be preset and displayed according to the marker position [10-11]. However, the mobile augmented reality map is oriented to objectives of different types and subjects, so it is necessary to consider multi-attribute parsing of markers, spatial information retrieval, virtual model construction, and model display adjustment in the process of fusion.

This study takes the mobile augmented reality map as the research object and designs and proposes a new user-defined nested hierarchical semantic marker in the map space. The authors study the hierarchical recognition and the analysis of spatial marker in the planar map based on augmented reality technology, 3D tracking registration, and virtual-real fusion methods and strive to establish a new geographic information visualization interface paradigm and application mode. The mobile augmented reality map enhances the multidimensional dynamic visual description of the real world, realizes the virtual-real fusion and enhanced expression in urban plane map, and provides a rich human-computer interaction mode for map application.

The main contributions of this study are as follows:

(1) This study proves the feasibility of using hierarchical semantic markers to implement mobile augmented reality map applications. Most of the markers in previous studies were just tags that replaced virtual models.

(2) The authors design and propose a new user-defined nested hierarchical semantic marker in the map space, which integrates AR mark and QR code information. The user-defined hierarchical semantic marker in the map space changes the fixed markup pattern in traditional augmented reality and can represent the spatial position, semantic information, and multidimensional attribute information of map elements with flexible features and universality. In addition, the marker can be used to assist the virtual-real fusion and model matching and lays a foundation for augmented reality map applications.

(3) The virtual-real fusion algorithm of a mobile augmented reality map and virtual model is proposed. The algorithm realizes hierarchical recognition of the urban map self-defined semantic marker, 3D tracking registration, and seamless integration of real scenes and virtual information.

(4) Through experiments, such as map marker level recognition and map virtual-real information fusion, and results analysis, it is shown that the user-defined marker and map model virtual-real fusion algorithm proposed in this paper is effective. The marker information is rich, the model is displayed accurately, and the effect of enhancing expression is achieved. Through an analysis of the recognition efficiency and recognition rate of the user-defined marker, it is proved that the recognition efficiency of the marker hierarchy recognition method proposed in this paper meets the real-time requirements of augmented reality; moreover, the recognition accuracy of the proposed method is high, and the method has a degree of universality.

2 A User-defined Hierarchical Semantic Marker in Urban Map Space

According to the spatial characteristics of various elements in the urban planar map, a user-defined hierarchical semantic marker for urban planar map space is designed based on a QR Code. Integrating AR mark and QR code information, the marker can represent the identification number, identification type, spatial location, and semantic information of map features and assist in tracking registration and virtual-real fusion to lay the foundation for enhanced expression of the urban planar map.

2.1 Characteristic Analysis of Map Symbols

Map symbols and their composition. Maps, which follow specific mathematical rules, use map language (that is symbolic system) to express the spatial distribution, combination, connection, change, and development of various things on the earth or other stars on a certain carrier through cartographic synthesis [12]. Maps, with the features of formulation, abstraction, and symbolization, use map languages such as points, lines, polygons, colors, and characters to express complex natural or social phenomena.

Maps have a specific symbol system, strict mathematical rules, and scientific cartographic generalization, including important information such as symbols, colors, coordinate systems, and characters [13]. The map symbol, as a basic means of expressing the content of the map, is closely related to the specific features of the geographical things in reality and consists of graphics and texts with different shapes, sizes, and colors. According to the geometric features and spatial distribution characteristics of the cartographic objects, the map symbols can be divided into point, line, face, and body symbols. These symbols can not only display the spatial position, range, quantity, quality, and attribute characteristics of the geographical objects, but also intuitively express the spatial distribution law, mutual connection, and dynamic changes of geographical objects.

The user-defined hierarchical semantic marker studied in this paper is aimed at the point and face features, and can fully express the spatial and attribute characteristics of these two features in the marker with individual characteristics.

Map visualization form. Taking geographic information visualization as its goal, the map focuses on the expression, processing, and transmission of geographic information. As carriers of geographic information, traditional maps mainly include ordinary maps, thematic maps, topographic maps, and image maps. With the continuous expansion of modern map concepts, new map forms have emerged one after another, such as augmented reality maps, electronic maps, digital maps, panoramic maps, and holographic maps.

With regard to the form of visualization, the map has undergone a transformation process from a two-dimensional to three-dimensional and even a multidimensional form, from a static to dynamic form, from a single direction to all-round form, from a paper media to multimedia form, and from reality and virtual forms to a virtual–real fusion form [14-15]. Among them, the augmented reality map constructed by augmented reality and visualization technology, enhanced expression of the morphological structure and attribute information of map space elements, and realization of virtual-real fusion are the current research hotspots [1-2, 16].

2.2 Design of the User-defined Hierarchical Semantic Marker in Map Space

In traditional augmented reality applications, it is necessary to use traditional AR identification or QR codes to complete rapid recognition and tracking registration of specific targets in a real scene. ARtoolkit uses a predefined specific graphical pattern that is a special visual tag. Fig. 1 shows a typical traditional Artoolkit augmented reality tag, which is usually square and has a black continuous boundary with a contrasting color background. The boundary content area contains special images. The traditional Artoolkit tag tracking registration effect is stable, but it is not suitable for the identification of massive tags [17]; furthermore, as it usually needs to save a limited number of tag images in mobile terminals, it is not suitable for multi-user augmented reality systems.

Fig. 1. Typical traditional Artoolkit augmented reality tag

The QR code is a matrix two-dimensional code that completes encoding in a rectangular space by using different distributions of black and white pixels in the matrix. At each position of the matrix element, a point with a specific shape is given a binary value of "1"; otherwise, it is given a binary value of "0." The arrangement of points determines the meaning represented by the QR code [18-19]. As shown in Fig. 2, the QR code includes position detection patterns, positioning patterns, correction patterns, format information, and data areas of three corner points in the area.

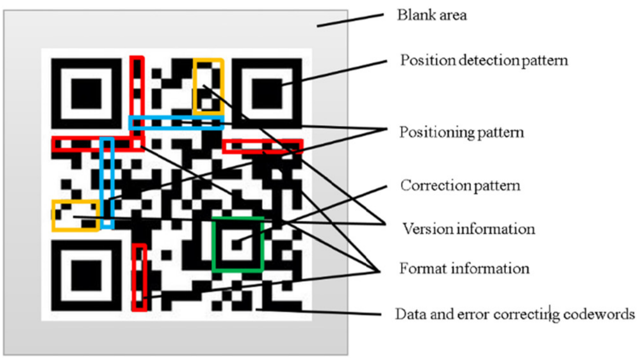

Fig. 2. Internal structure diagram of a traditional QR code

According to the expression requirements of the characteristics of city map elements (spatial point features, surface features), we propose the user-defined nested hierarchical semantic marker in map space, which combines AR mark and QR code information. As shown in Fig. 3, the AR marker has a black border that can be divided into a 7 × 7 grid with an internal 5 × 5 grid containing a specific ID code. The black grid in the grid represents 0 and the white grid represents 1. The augmented reality marker can not only represent the spatial location of map elements, but also provide information such as map feature type, feature identification, and attribute description, and can be used for virtual-real fusion and model matching. It mainly includes an identification number, identification name, identification type, virtual model number, and spatial position coordinates. An image recognition method based on computer vision is used to detect and identify the pre-added user-defined hierarchical semantic marker in map space and to obtain the map feature type and its auxiliary information, which lays the foundation for generating virtual model and virtual-real fusion.

Fig. 3. User-defined hierarchical semantic marker in map space

3 Hierarchical Semantic Marker Recognition on Urban Maps and 3D Tracking Registration

3D tracking registration is key to achieving virtual-real fusion between maps and models [20]. Based on the hierarchical semantic marker recognition on urban maps, the virtual model, detailed attribute description, etc. are obtained. The positioning process of a virtual model in the real world is called 3D registration. Before 3D registration, you need to obtain the position, angle, and motion of the camera (i.e., observer's perspective), and unify the different coordinate systems. This process is called tracking, and it can be achieved through sensors or the computer vision method. The latter method is adopted in this study.

3.1 User-defined Marker Recognition and Registration

In marker-based augmented reality applications, recognition of markers in a video is the first step of the virtual-real fusion algorithm. Based on the user-defined marker integrating spatial markers and QR code information adopted in this study, a hierarchical recognition method is used for recognition. In this method, QR code data are first acquired by using a two-dimensional code image extraction algorithm, and the spatial position and description information of markers contained therein are obtained by QR code decoding. Then, image recognition is carried out for traditional augmented reality spatial markers, and their spatial position in a plane map is recorded, which lays a foundation for completing the three-dimensional registration of virtual-real models.

The recognition process of the user-defined hierarchical semantic marker is shown in Fig. 4. First, recognition in the QR layer is performed. Image frames are captured by a camera and then graying processing is performed on the frames. De-noising processing is performed by Gauss smoothing filtering and mean filtering. After binary processing is performed on the image, closed operation, corrosion, expansion, and other methods are used to calculate and obtain the rectangular boundary of the QR code. Through angle correction and distortion correction, the position detection pattern is tracked and the QR code is identified and semantically interpreted.

Based on identification in the QR layer, recognition in the AR layer is performed. After obtaining the AR markers, their modes were mapped according to their types, and data such as the camera position and the line of sight were calculated to finally realize the three-dimensional tracking, which lays a foundation for three-dimensional registration of the virtual model. The process is shown in Fig. 4.

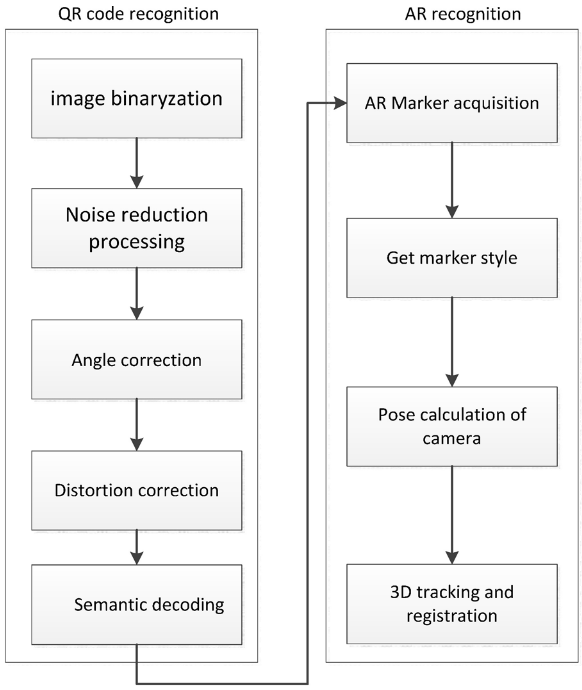

Fig. 4. Flow chart of user-defined semantic marker hierarchical recognition

3.2 3D Tracking Registration Method

Real-time tracking registration. The registration process of realistic scenes and virtual models in three-dimensional space is called three-dimensional registration [21]. The computer vision-based registration method is currently a mainstream method. The 3D registration of virtual models mainly involves four coordinate systems: a world coordinate system, virtual space coordinate system, camera coordinate system, and screen plane coordinate system [22]. As shown in Fig. 5, the purpose of three-dimensional registration is to bring these four types of coordinate systems together to achieve a seamless integration of real-world scenes and virtual information.

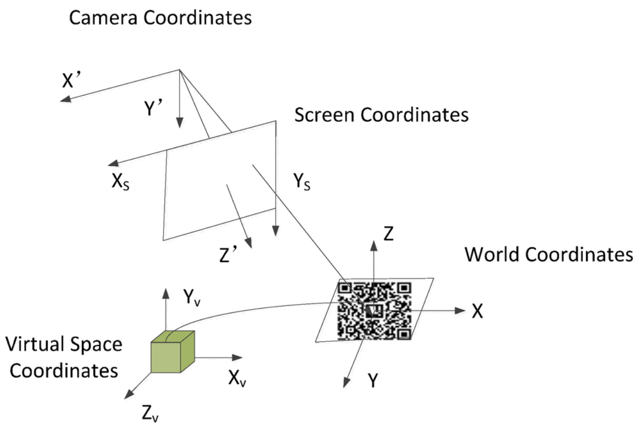

Fig. 5. Principles of 3D tracking registration

Calculation method for 3D tracking registration. Augmented reality systems mainly adopt two kinds of tracking registration technologies: hardware device-based tracking registration technology and computer vision-based tracking registration technology. The latter was adopted in this study. Computer vision-based tracking registration is used to determine the relative position and orientation of the camera and the target in the real world by giving one or more images. This study uses the affine transformation to realize three-dimensional registration.

The space affine transformation in computer vision needs to calculate projection coordinates from virtual objects to real scene images, i.e., under homogeneous coordinate representation, the conversion relationship from a point on virtual objects in a three-dimensional space to its corresponding projection point on the real scene images is established. By constructing a global affine coordinate system, the conversion relationship between the virtual space coordinate system and the real space coordinate system can be obtained. Then, the camera coordinate system and the image plane coordinate system can be established; furthermore, the virtual model can be three dimensionally registered according to the camera position and posture data [23-24].

It is assumed that the real space coordinate system is $[x, y, z, 1]^T$, the virtual space coordinate system is $[xv, yv, zv, 1]^T$, and the camera space coordinate system is $[x', y', z', 1]^T$, wherein the conversion relationship between the real space coordinate system and the virtual space coordinate system is

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = U_{4\times4} \begin{bmatrix} x_v \\ y_v \\ z_v \\ 1 \end{bmatrix}. \tag{1}$$

The conversion relationship between the camera space coordinate system, the world coordinate system, and the virtual space coordinate system is

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = V_{4\times4} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = V_{4\times4} U_{4\times4} \begin{bmatrix} x_v \\ y_v \\ z_v \\ 1 \end{bmatrix}, \tag{2}$$

where $V_{4\times4}$ in formula (2) represents a transformation matrix between the camera space coordinate system and the world coordinate system, which can be realized by translation and rotation. In other words, the coordinate system transformation is completed by combining the rotation matrix $R_{3\times3}$ and the translation matrix $T_{3\times1}$, and their parameters can be calculated by the position, direction, and attitude of virtual objects with respect to the camera space coordinate system. As the relationship between the world coordinate system and the model virtual space coordinate system has been known, solving $V_{4\times4}$ is key and can be computed by

$$V_{4\times4} = \begin{bmatrix} R_{3\times3} & T_{3\times1} \\ 0 & 1 \end{bmatrix}. \tag{3}$$

Formula (4) can be derived from Eqs. (2) and (3), and is defined as

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} R_{3\times3} & T_{3\times1} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}. \tag{4}$$

Therefore, $[x', y', z']^T$ can be defined as

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} R11 & R12 & R13 \\ R21 & R22 & R23 \\ R31 & R32 & R33 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} t1 \\ t2 \\ t3 \end{bmatrix}. \tag{5}$$

As shown in

$$T_{3\times1} = \begin{bmatrix} t1 \\ t2 \\ t3 \end{bmatrix}. \tag{6}$$

and

$$R_{3\times 3} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} \cos\omega & 0 & -\sin\omega \\ 0 & 1 & 0 \\ \sin\omega & 0 & \cos\omega \end{bmatrix}, \quad \textbf{(7)}$$

where the translation matrix $T_{3\times 1}$ shifts the coordinate system origin of the camera to the origin of the world coordinate system so that both origins coincide with each other. The rotation matrix $R_{3\times 3}$ is a rotation matrix when the camera coordinate system is rotated to coincide with the world coordinate system, and it can be regarded as a result of respective rotations around the x, y, and z axes.

4 The Virtual-real Fusion Algorithm of Map Models

On the basis of augmented reality technology, three-dimensional enhanced expression of typical spatial elements in urban maps needs to use terminals to interact naturally with the actual environment to achieve seamless integration of virtual models and real space. Therefore, virtual models and information should be superimposed directly onto actual scenes [25]. The virtual-real fusion needs to consider the occlusion relationship, display rationality, and illumination consistency between virtual models and the real world, etc., and set the position, angle, and size of the model display to enhance the expression effect.

The virtual-real fusion process of map models in this paper mainly involves hierarchy recognition and tracking of map markers, three-dimensional registration, spatial information retrieval, virtual model construction, fusion display, and other steps. The overall algorithm flow is shown in Fig. 6.

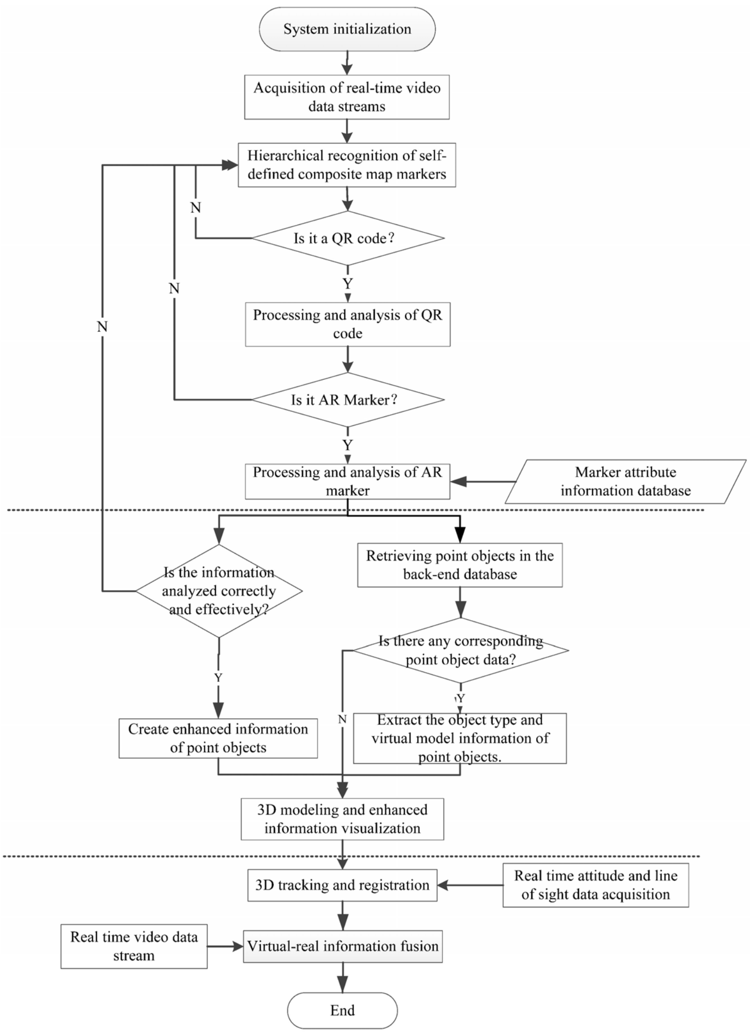

Fig. 6. Virtual-real fusion algorithm flow of the map models

The virtual-real fusion algorithm of map models can be divided into three parts. The first part mainly includes hierarchical recognition of composite map markers based on real-time video data streams. By identifying and parsing the user-defined map markers step by step, the acquisition of initial point feature attribute information can be completed. In the second part, the parsing results of user-defined composite map markers are used to retrieve and extract point map feature ID, object type, and virtual model information in the background database, generate point-like feature enhancement information and fuse the above data for 3D modeling, and enhance information visualization. In the third part, the real-time position, attitude, and line-of-sight data, together with real-time video data streams, are utilized for 3D tracking and registration to complete fusion of virtual-real models and information.

5 Virtual-real Fusion Experiment and Analysis of the Augmented Reality Map

5.1 Virtual-real Fusion Experiment Method

Based on the foregoing research, an augmented reality map virtual-real fusion system, ARMapV1.5, was developed. This system realizes the functions of city map marker recognition, tracking, and virtual-real fusion. The ARMap system uses a ThinkPad P51 workstation with a CPU Intel(R) Core(TM) 7820HQ and16G memory as the hardware platform, and it uses a high-definition camera (5 million pixels, resolution 1920 × 1080) to obtain realistic images. Windows 10 was adopted as the operating system, Visual C++2015 was adopted as the system development tool, ARToolkit was used as the augmented reality development kit, OpenCV2.4.11 was used as the computer vision development kit, and OpenGL 4.0 was used for 3D model modeling and visualization.

The ARMap system consists primarily of a real-time data acquisition module, a marker detection and recognition module, a real-time tracking and registration module, and a virtual-real fusion module. This system can detect and analyze real-time images collected by the camera, complete the hierarchical recognition of markers and analysis of auxiliary information, realize adaptation and registration functions of the virtual model with the real world, and achieve real-time tracking from the observer's field of view and 3D registration of virtual models. Furthermore, by integrating 3D modeling and the virtual reality, the virtual-real fusion of map models is realized in the real world.

The Beijing urban area map was taken as an example, and Beijing Information Science and Technology University and other places were selected as the experimental area. The ARMap system was used to verify the correctness and effectiveness of the models, methods, and results in the study. The virtual-real fusion experiment of augmented reality map mainly includes hierarchical recognition and information parsing of user-defined composite map markers, 3D tracking and registration of the virtual models, and virtual-real fusion of map models.

Calibrating the camera is the first step in the process of augmented reality virtual-real fusion. For a specific camera, a separate parameter file can be generated by the camera calibration technique to obtain a camera's intrinsic matrix and to ensure a one-to-one congruent relationship between the 3D space and 2D screen in the camera model.

5.2 Hierarchy Recognition of Map Marker and Information Analysis

In the course of marker recognition in the QR layer, the QR code image extraction algorithm is used to obtain and recognize the QR code, and acquire location and description information hidden in the spot object marker by QR code analysis. As shown in Fig. 7(a), the blue square frame in the figure represents the image in the QR layer from the map space user-defined markers automatically recognized by the system. Through QR Code analysis, the information in this marker is "B0121, Teaching building 1, Beijing Information Science and Technology University, Point, 116.38826, 39.99528."

As shown in Fig. 7(b), the information parsing in the AR layer maintained its spatial position in the plane map by identifying the traditional AR space marker, and it combines the identification position and description information obtained by decoding the QR code. It lays the foundation for the information fusion of the virtual-real map model in the plane map.

(a) Effect of 2D code recognition and analysis

(b) Effect of point feature recognition and information analysis

Fig. 7. Hierarchy recognition and information analysis of point features

5.3 Virtual-real Information Fusion of an Urban Map

After hierarchy recognition and analysis of the spatial markers for point features, the position of the marker in a plane map is recorded and the description information and model type obtained by decoding the QR code are gained. The query is performed in a terminal database table according to the point feature ID and the model type, and then the feature type, virtual model ID, model data, and the like are obtained to generate the augmented information of point features. Using a camera's line of sight and view frustum data, 3D registration is achieved by back-calculating the actual spatial position of the corresponding model with markers. By combining the above data and location for 3D modeling and visualization, and adjusting the position and posture of the 3D virtual models to match the observer's perspective, virtual-real fusion can be achieved.

Fig. 8(a) is the virtual-real fusion effect diagram with a single map marker, in which the marker can be identified as "the first building of Beijing Information Science and Technology University." The MARMap system program is used to detect, analyze, and obtain enhanced information, which it then displays on the terminal to realize virtual-real fusion of the models and the markers.

Fig. 8(b) is the virtual-real fusion effect diagram with multi maps markers, in which there are three typical symbols, including Beijing Information Science and Technology University and Directorate of Imperial Academy and Zizhuyuan Park. The MARMap system program can simultaneously detect, identify, and analyze multiple markers. Moreover, according to the corresponding models preset in the database, 3D modeling and visualization are carried out; finally, the virtual-real fusion is realized by adjusting the position and posture of the models.

(a)Virtual-real fusion effect diagram with a single map marker

(b) Virtual-real fusion effect diagram with multiple map markers

Fig. 8. Virtual-real fusion effect diagram

5.4 Analysis of Experimental Results

Effectiveness analysis of user-defined marker recognition and virtual-real fusion. The experiment focuses on objects with different locations or different types, such as tourist attractions, schools, and office buildings, and each object corresponds to a different custom marker and model.

The results of the virtual-real fusion experiment for the map model with the single marker and the multi markers, reveal that a custom nested map marker combining AR marker with QR code information can effectively express characteristics of urban map features and provide information such as map feature type, spatial location, and attribute description.

The custom marker is recognized hierarchically based on the computer vision method. In such a method, map feature type and model information can be obtained by analysis, and the real-time locations, sights, and other data can be used for 3D tracking and registration. By combining this information with real-time video data streams, the virtual-real model and information fusion are achieved.

Hierarchy recognition efficiency and rate analysis of the custom markers. To prove the hierarchy recognition efficiency and rate of the custom markers in this paper, the virtual-real fusion experiments of the map model with a single marker were performed under the background and conditions described above.

The experimental results are shown in Table 1. Three groups of experiments were performed. The recognition frequencies of the three groups were 25 Hz, 100 Hz, and 200 Hz. Each group of experiments was performed three times, and the data in Table 1 is the average value of three experiments. Three parameters, i.e., the single mean time for QR code recognition (t1), single mean time for hierarchical recognition (t2), and accuracy of recognition, were recorded in each group of experiments when the number of recognition (N) was 200, 400, 600, 800, 1000, and 1200.

It can be drawn from Table 1 that t1 was 58.820 ms, 57.287 ms, and 55.412 ms at 25 Hz, 100 Hz, and 200 Hz, respectively. Furthermore, at recognition frequencies of 25 Hz, 100 Hz, and 200 Hz, t2 was 80.473 ms, 78.467ms, and 76.618 ms, respectively.

From the overall data, the average t1 and t2 values were 57.173 ms and 78.519 ms, respectively. The recognition speed of the markers was stable. However, as the recognition frequency increased, the recognition time decreased. In addition, the smaller the recognition frequency, that is, the larger the gap between two recognition methods, the higher is the recognition accuracy rate.

This experiment shows that the hierarchy recognition method for markers is fast. Furthermore, the recognition efficiency meets the real-time requirement of augmented reality, and the recognition accuracy is quite high.

Table 1. Custom recognition efficiency table (frequency 200 Hz)

Team NO.	Recognition frequency (Hz)	Recognition times N	Single mean time for QR code recognition t1 (ms)	Single mean time for hierarchical recognition t2 (ms)	Recognition accuracy rate (%)
1	≤25	200	58.352	79.553	100
		400	58.737	80.138	100
		600	58.897	80.515	100
		800	58.911	80.672	100
		1000	58.991	80.895	100
		1200	59.033	81.063	100
2	≤100	200	54.88	74.913	100
		400	56.928	77.753	100
		600	57.54	78.755	100
		800	57.906	79.404	100
		1000	58.143	79.818	100
		1200	58.326	80.158	100
3	≤200	200	50.273	70.258	99.805
		400	54.559	75.437	99.91
		600	56.011	77.29	99.942
		800	56.811	78.318	99.957
		1000	57.261	78.965	99.966
		1200	57.554	79.438	99.972

Experiment summary. The experimental results reveal that the custom nested map marker changes the fixed marker mode in traditional augmented reality and adds the multi-dimensional attribute information and spatial information of the markers, making the markers more flexible. The virtual-real fusion algorithm of the map models can seamlessly integrate the virtual model corresponding to the marker with the map in reality, and the model display is accurate and stable. The effect of enhanced expression is achieved.

Through the validity and efficiency verification experiments of the hierarchy recognition of the markers, it is illustrated that the user-defined nested map marker proposed in this paper is effective, as it has a high recognition efficiency that meets real-time requirements and a high accuracy rate.

Other experiments were performed on different map media, such as paper printed maps and electronic maps as well as different colored maps, all of which achieved good results, indicating that the proposed method has a degree of universality.

6 Conclusion

Augmented reality technology is applied to enhance expression of urban planar maps in this paper. Based on the characteristic analysis of map symbols, a custom hierarchical semantic marker in map space combining AR marker and QR code information is designed and proposed. The marker can represent spatial locations and attribute information of map elements and assist virtual-real fusion. The hierarchy recognition and tracking method and process of the urban map marker are introduced in detail, and the virtual-real fusion algorithm of the map model is proposed.

The analysis and experimental results reveal that the user-defined marker and virtual-real fusion of the map model studied in this paper are effective. The hierarchical recognition efficiency and recognition accuracy rate analysis of the custom map markers further reveal that the proposed method is fast; moreover, its recognition efficiency meets the real-time requirements of augmented reality, and its recognition accuracy is higher.

On the basis of the findings of this study, further studies on virtual-real fusion of mobile augmented reality maps based on mobile terminals are needed. In future research, high-precision sensor modules, such as magnetometers and gyroscopes, could be integrated into the MARMap system. In the MARMap V2.0 system, a high-definition camera could be used to obtain realistic images; furthermore, a magnetometer and gyroscope could be used to obtain the camera direction and attitude data to achieve a more accurate virtual-real fusion effect.

Acknowledgements

This study is supported by The Scientific Research Project of Beijing Educational Committee (KM201811232010), Beijing Science and Technology Innovation Service Capacity Building-Fundamental Research Fund (PXM2019_014224_000026), The Project for Promoting Connotative Development of Universities-"Information +" project, Open Research Funding Program of KLGIS (KLGIS2015A02). We would like to thank the anonymous reviewers for their constructive comments that greatly improved the quality of the paper.

References

[1] P. Volker, S. Monika, Augmented paper maps Exploring the design space of a mixed reality system, International Journal of Photogrammetry and Remote Sensing 65(3)(2010) 256-265.

[2] J. Bobrich, S. Otto, Augmented Maps, in: Proc. International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 2002.

[3] A. Morrison, A. Mulloni, S. Lemmelä, A. Oulasvirta, G. Jacucci, P. Peltonen, D. Schmalstieg, H. Regenbrecht, Collaborative use of mobile augmented reality with paper maps, Computers & Graphics 35(4)(2011) 789-799.

[4] A.D. Lucia, R. Francese, I. Passero, G. Tortora, Augmented reality mobile applications, challenges and solutions, Recent Patents on Computer Science 4(2)(2011) 80-90.

[5] M. Sun, X. Chen, F. Zhang, H. Zhang, Augmented reality geographic information system, Acta Scientiarum Naturalium Universitatis Pekinensis 40(6)(2004) 906-913.

[6] H. Uchiyama, H. Saito, M. Servieres, AR city representation system based on map recognition using topological information, Virtual and Mixed Reality, Lecture Notes in Computer Science 5622(2009) 128-135.

[7] J. Chen, Y. Wang, Q. Shi, D. Yan, Dynamic registration of augmented reality systems, Transactions of Beijing Institute of Technology 5(2001) 631-636.

[8] D. Schmalstieg, G. Reitmayr, Augmented reality as a medium for cartography, In: W. Cartwright, M.P. Peterson, G. Georg (Eds.), Multimedia Cartography, Springer, New York, 2006, pp. 267-282.

[9] X. Wu, F. Ren, Q. Du, Affine transformation virtual and real matching spatial information enhanced expression, Computer Engineering and Applications 3(2010) 16-19.

[10] X. Fang, Q. Qu, Paper map expression and its application based on mobile augmented reality, Microcomputer & Its Applications 33(7)(2014) 41-43+47.

[11] A. Zhang, J. Zhuang, Q. Qi, Visualization and interaction of augmented paper maps based on augmented reality, Tropical Geography 32(5)(2012) 476-480.

[12] K. Liao, Modern Cartography, Science Press, Beijing, China, 2003.

[13] E. Garbin, F.D. Santil, J.V. Bravo, Semiotics and the cartographic visualization theory: considerations on the analysis of map design, Bol. Ciênc. Geod 18(4)(2012) 624-642.

[14] W. Dong, J. Wang, Effectiveness and efficiency of map symbols for dynamic geographic information visualization, American Cartographer 39(2)(2012) 98-106.

[15] R. Qin, N. Ye, H. Xu, Visualization of multi-dimension oceanographic data in geography information system, Journal of Tongji University 37(2)(2009) 272-276.

[16] R. Ford, R.K. Westhead, M. Smith, Mobile spatial mapping and augmented reality applications for environmental geoscience, International Journal of Internet Technology & Secured Transactions 2(1-4)(2013) 185-190.

[17] X. Lei, Z. Wu, J. Yong, Artoolkit 3D registration based on dynamic feature template matching, Computer Applications and Software 34(7)(2017) 104-109.

[18] T. Kan, C. Teng, W. Chou, Applying QR code in augmented reality applications, in: Proc. International Conference on Virtual Reality Continuum and ITS Applications in Industry, 2009.

[19] N. Park, W. Lee, W. Woo, Barcode-assisted planar object tracking method for mobile augmented reality, in: Proc. 2011 IEEE International Symposium on Ubiquitous Virtual Reality, 2011.

[20] S. You, U. Neumann, Fusion of vision and gyro tracking for robust augmented reality registration, in: Proc. 2001 IEEE Computer Society Virtual Reality Conference, 2001.

[21] J. Rekimoto, Matrix, A. Realtime, Object identification and registration method for augmented reality, in: Proc. 2002 IEEE Asia Pacific Computer Human Interaction, 2002.

[22] Q. Shi, Y. Wang, J. Cheng, Vision-based algorithm for augmented reality registration, Journal of Image & Graphics 7(7)(2002) 679-683.

[23] J.J. Koenderink, A.J.V. Doorn, Affine structure from motion, Journal of the Optical Society of America 8(2)(1991) 377-385.

[24] X. Wu, F. Ren, Spatial information augmented representation using affine transformations registration, Computer Engineering & Applications 46(3)(2010) 16-19+29.

[25] Y. Park, W. Woo, The ARTable: An AR based tangible user interface system, In: Z. Pan, R. Aylett, H. Diener, X. Jin, S. Göbel, L. Li (Eds.), Technologies for E-Learning and Digital Entertainment, Springer, Berlin, Heidelberg, 2006, pp. 1198-1207.