3D RECONSTRUCTION OF CULTURAL TOURISM ATTRACTIONS FROM INDOOR TO OUTDOOR BASED ON PORTABLE FOUR-CAMERA STEREO VISION SYSTEM

Zhenfeng Shao [a], Congmin Li [a,*], Sidong Zhong [b], Bo Liu [c], Honggang Jiang [c] , Xuehu Wen[d]

[a] State key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, 430079, China
[b] Wuhan University, Wuhan, 430079, China
[c] Geomatics Center of Guangxi, 5 Jianzheng Road, Nanning China,530023
[d] The Third Surveying and Mapping Engineering Institute of Sichuan, Chengdu Sichuan ,610500
shaozhenfeng@whu.edu.cn, cminlee@whu.edu.cn, sdzhong@whu.edu.cn,676802836@qq.com, 19743408@qq.com, 398169068@qq.com

*cminlee@whu.edu.cn

KEY WORDS: Cultural Tourism Resources Protection, Building Information Modelling, Indoor-outdoor Seamless Modelling, Four-camera Stereo Photographic Measurement System, 3D Reconstruction, Image-based reconstruction

ABSTRACT:

Building the fine 3D model from outdoor to indoor is becoming a necessity for protecting the cultural tourism resources. However, the existing 3D modelling technologies mainly focus on outdoor areas. Actually, a 3D model should contain detailed descriptions of both its appearance and its internal structure, including architectural components. In this paper, a portable four-camera stereo photographic measurement system is developed, which can provide a professional solution for fast 3D data acquisition, processing, integration, reconstruction and visualization. Given a specific scene or object, it can directly collect physical geometric information such as positions, sizes and shapes of an object or a scene, as well as physical property information such as the materials and textures. On the basis of the information, 3D model can be automatically constructed. The system has been applied to the indoor-outdoor seamless modelling of distinctive architecture existing in two typical cultural tourism zones, that is, Tibetan and Qiang ethnic minority villages in Sichuan Jiuzhaigou Scenic Area and Tujia ethnic minority villages in Hubei Shennongjia Nature Reserve, providing a new method and platform for protection of minority cultural characteristics, 3D reconstruction and cultural tourism.

1. INTRODUCTION

With the rapid development of Chinese tourism, cultural tourism resources become a new direction. It is widely believed that the development of cultural tourism resources brings prosperity to local social, economic and cultural, while it may cause changes in local ethnic culture. Traditional cultural in some areas is dramatically changing in clothing, architecture, customs and the way of life, which are rapidly converging with the foreign culture. In fact, local culture is facing the danger of losing individuality and characteristics. For example, in the recent years, new residential construction in Jiuzhaigou are mainly made of brick or reinforced concrete structure, while traditional cedar chips roof is replaced by tile roof, and rammed earth walls were changed to brick. This type of building has a richer exterior decoration and colour and integrates more Han Chinese architectural elements. Actually, these changes enhance the comfort and convenience of living in Tibetan villages. However, the unique architectural culture of minorities is disappearing. In order to alleviate the contradiction between cultural resource protection and utilization in cultural tourism, three-dimensional reconstruction technologies have become an important means to protect and present cultural tourism resources.

At present, the common used method of three-dimensional reconstruction of a scene or object is mainly depended on existing interactive modelling software. Although the modelling accuracy can meet the requirements of practical applications, it needs heavy manual operation, and thus has lower efficiency in modelling and higher cost. Among the automatic 3D reconstruction techniques, laser scanners and image-based approach have attracted widespread attention. The former is widely used in the digitization of cultural heritage, because it is able to quickly scan the surface of the target area, and generate high-density 3D point cloud data. But these devices are relatively expensive and are generally used for surface modelling of buildings. While the latter has lower hardware requirement and can construct models at lower cost. Moreover, 3D models obtained by the latter can basically meet certain requirements. A real 3D model should contain not only the description of the appearance of the building, but also its internal structure. As respect of data acquisition, modelling for indoor lacks of effective 3D data acquisition equipment at present.

Based on the above requirements and analysis, a portable four-camera stereo vision system is developed and implemented, which has been successfully applied to 3D reconstruction of architectural features from indoor to outdoor in two typical cultural tourism resources in China. The system can directly

---

* Corresponding author

measure the geometry and attribute information by the acquired images, and automatically model based on this information. It solves the problems of indoor-outdoor data acquisition and integrated modelling, and provides a new method and platform for the protection of minority cultural characteristics, 3D reconstruction and cultural tourism.

2. RELATED WORK

Compared with 3D range scanning, Image-based 3D reconstruction techniques have been considered as a low-cost and effective tool for producing high-quality 3D models of real world in terms of hardware requirements, knowledge backgrounds and man-hours (Koutsoudis, et al, 2014). Researchers have done considerable work to create high-quality 3D models, especially in computer vision community. Among these techniques, the Structure-From-Motion (SFM) (Robertson& Cipolla, 2009) and Dense Multi-View 3D Reconstruction (DMVR) (Furukawa& Ponce, 2010) are the most popular algorithms. These methods can be used to recover a scene or object by processing numerous unordered images taken from perspective viewpoints with large overlapping areas. The pipeline of image-based 3D reconstruction is generally composed of different steps that can be automatically performed. First, camera parameters can be obtained by matching corresponding features or calibrated in a control field. According to the camera imaging model, 3D geometry information as well as texture can be restored. A number of software based on the two algorithms has been made available, such as Bundler (Snavely, 2008), Visual SFM(Wu, 2011), etc.. Based on these basic theories, image-based 3D reconstruction techniques have been widely used. However, these applications mainly focus on surface modelling of a scene or building, which can hardly meet the requirements of fine interior modelling.

In the research of stereo vision, according to the number of cameras, stereo vision measurement methods can be mainly divided into three categories: monocular stereo vision measurement method, binocular stereo vision measurement method and multi-view stereo vision measurement method. Due to self-occlusion and shadows of the measured objects, as well as the limited field of view of a camera and depth of field, neither monocular nor binocular stereo measurement system can get the whole surface data of the measured object through a single measurement. Usually, it needs to increase the times of measurement from a different angle range. When the number of cameras reaches certain level, it will add the difficulty of subsequent processing, which can be a big challenge for data mosaic and integration. Therefore, the developed portable four-camera measurement system can be of great significance.

3. PROTABLE FOUR-CAMERA BASED STEREO VISION SYSYTEM

Four-camera Stereo photogrammetry system is a type of image-based non-contact measurement. Firstly, it obtains four images of the same object or scene simultaneously from four different perspectives by using four relatively fixed cameras. Then, measurements can be carried out on these images, thus, physical geometric information such as 3D description of the object coordinates, sizes and shapes of an object or a scene, as well as physical property information such as the materials and textures can be automatically achieved. On the basis of the information, 3D model can be automatically constructed. Section 3.1 and 3.2 respectively introduces the hardware design and methods of 3D model acquisition of the developed system.

3.1 Hardware Platform

As illustrated in figure 1, the system is mainly composed of four digital cameras, rigid connection rod and synchronous controller. Those four cameras are divided into two groups on average and each group is fixed at ends of the rigid connecting rod within the camera cover, thus, the relative position and orientation of the camera can be maintained during shooting. The most important part of the measurement system is the synchronous controller. It is capable of receiving signals from a handheld remote shooting and then transmits high precision synchronization signals to cameras, ensuring synchronization of four cameras shooting.
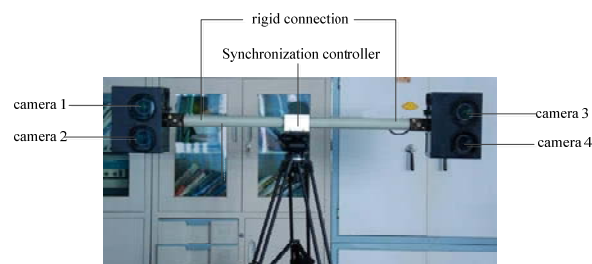

Figure 1. Hardware structure of portable four-camera stereo vision system

3.2 Methodology of fully automatic 3D models acquisition

Stereoscopic image collected by the four cameras can be transmitted to computers in real-time. When measuring, three-dimensional coordinates of the target can be obtained by the computer automatically calculate and applied directly to the object three-dimensional positioning, geometric measurement and object reconstruction. Procedures of acquiring 3D models through the proposed system are described as Figure 2.
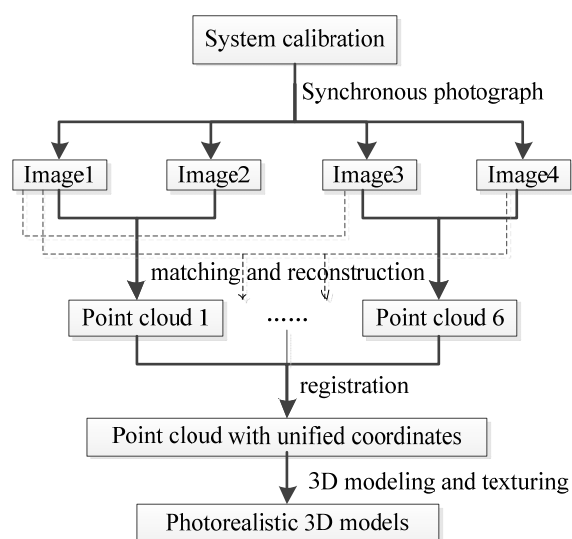

Figure 2. General pipeline for delivering photorealistic 3D models from collections of portable four-camera stereo vision system.

3.2.1 Camera calibration: Calibrating this system to obtain camera parameters is one of prerequisite for the measurement. Because of the portability of the system, only digital cameras or digital video cameras are needed, not requiring professional cameras. However, considering the accuracy of cameras and precision of 3D reconstruction, the camera calibration of the system is necessary. Images of the control field acquired by four cameras existing in this system can be utilized to calculate camera parameters, including internal and external orientation elements of cameras, and camera distortion parameters etc.. Thus, there is no need to providing rulers or setting known points. Once the system is calibrated in the professional calibration field, its parameters can be kept.

3.2.2 Data acquisition: When acquiring indoor-outdoor data, the special synchronous controller must be used to control shutter switch signal of four cameras, ensuring synchronization photography of four cameras to prevent camera ego motion and parameters failure caused by dynamic scenes and the movement of objects. In order to obtain high accuracy of texture data in the measured area as well as providing guidance for the automatic matching of the whole scene, UAVs will be applied to assist modelling.

3.2.3 Feature extraction and image matching: After acquiring data, a series of pre-processing operations for images are needed, such as feature extraction and stereo matching. In the paper, SIFT algorithm (Lowe, 2004) is used for image feature extraction because of its scale and rotation invariant characteristics. As for matching, matching strategy based on the principle of nuclear lines intersecting is proposed. Combining with camera imaging model, each feature point can be constructed into six imaging binocular stereo vision pairs when using four-camera stereo vision system. As is shown in figure 3, a reality scene point can be imaged on four pictures, assuming to be p_1, p_2, p_3, p_4, and l_{12}, l_{13}, l_{14} l_{23}, l_{24}, l_{34}, l_{14} respectively stands for the nuclear line of the six stereo pairs. When extracting the feature point p_1 in the upper left picture, it is easy to find the corresponding point p_3 in the lower left image with high accuracy. According to the epipolar constraint, nuclear line l_{12} and l_{23} can intersect at point $p_{2'}$, which is corresponding point of p_1 in the upper right image. Similarly, it is easy to find corresponding points p_2, p_4 and $p_{4'}$, which is the intersection point of the two nuclear lines l_{24}, l_{34}.

Therefore, whether the match of point p_2 and p_4 is correct can be verified by judging whether the two points p_4 and $p_{4'}$ coincide. If the two points do not coincide, it indicates a mismatch and needs to find another point until all of the matches are correct.
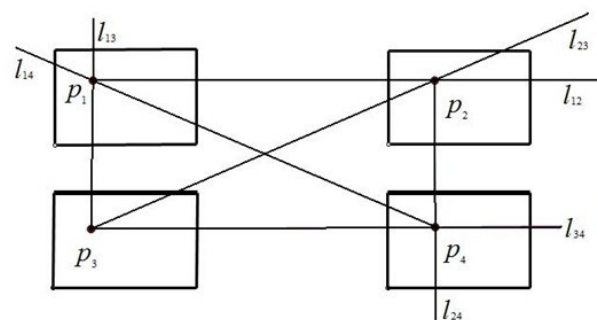

Figure 3. Corresponding points matching based on the principle of nuclear lines intersect.

The biggest advantages of the four-camera system is that it can not only greatly reduce the searching area for feature point matching, but also eliminate the gross error and providing redundant observation for checking and ensuring the accuracy of matching.

3.2.4 3D modelling and texturing: Once stereo images are transmitted over the network into the computer, the 3D spatial information can be automatic obtained by professional software. From 3D space information, 3D point cloud data and texture data can be extracted. Then 3D point cloud data can be used for automatic construction of triangulation. At last, finer precise 3D models will be built through dense matching. Detail implementation can refer to the open-source software bundler (Robertson& Cipolla, 2009).

3.2.5 Scene modelling and texturing: After extracting 3D models of indoor and outdoor areas, unmanned aerial vehicles images are used to assist scene modelling, such as helping for extracting top texture of architecture and providing guidance for the whole scene matching.

4. APPLICATIONS IN Reconstruction CULTURAL TOURISM RESOURCES

4.1 Experimental Area

The system is applied to the indoor-outdoor seamless modelling of distinctive architecture existing in two typical cultural tourism zones in China, that is, Tibetan and Qiang ethnic minority villages in Sichuan Jiuzhaigou Scenic Area and Tujia ethnic minority villages in Hubei Shennongjia Nature Reserve.

Jiuzhaigou Scenic Area is located on the edge of the Tibetan Himalayan Plateau in Northern Sichuan Province, which is named for nine Tibetan villages and it was declared a UNESCO World Heritage Site in 1992. Historically external blocking traffic, specific historical and closed conditions make Jiuzhaigou retain the original nature of traditional Tibetan culture. Building facilities in Jiuzhaigou Scenic Area are mainly nine Tibetan villages and temples. Among the nine villages, Zezawa Village, Shuzheng Village and Heye Village are well-preserved and have frequent tours.

Shennongjia Nature Reserve is located in the northwest of hubei province. In 1990, it was incorporated in the UNESCO World Network of Biosphere Reserves. People living in Shennongjia forest region are mainly Tujia ethnic minority. Diaojiaolou with

red lanterns and Tunkou have become a distinctive architectural culture.

For the unique cultural buildings and landscape in the two cultural tourism zones, 3D reconstruction work has been conducted by the developed system, so as to preserve the unique cultural tourism resources as much as possible. The workflow is described as figure 4.
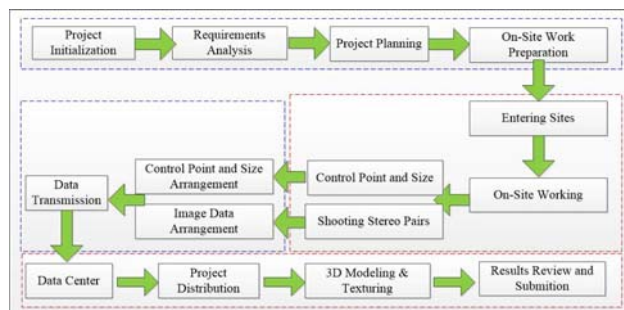

Figure 4. Workflow of 3D modelling on-site.

4.2 Results and Discussions

This study mainly focuses on the 3D reconstruction of the outdoor surface and indoor structure for two typical cultural tourism resources. Experiment results show that it is feasible to precede indoor-outdoor reconstruction with the use of the portable four-camera stereo photographic measurement system.

The measurement system is a type of image-based non-contact measurement. As for image collection, the specialised four-camera device can be manipulated by one person to obtain both facade and interior structure data of the buildings. And unmanned aerial vehicles images are used to help for extracting top texture of architecture. Once these images are collected, they will be transmitted to the computer to be measured through the Internet in real-time. Then 3D models can be automatically constructed by the professional software. Because of the distribution characteristics of the four-camera, image matching can be greatly efficient and accurate. Finally, 3D models containing detailed descriptions of both its appearance and its internal structure can be constructed. For the purpose of visualization of cultural tourism resources, 3D models can be presented and manipulate in a specialized software, meanwhile, they can be saved as AVI form. The final photorealistic 3D models of the distinctive architecture are shown in figure 5-8.

Figure 5. 3D reconstruction of Heye Village in Jiuzhaigou Scenic Area

Figure 6. Internal-outdoor reconstruction of residential building with Tibetan characteristics in Jiuzhaigou Scenic Area

Figure 7. 3D reconstruction of distinctive building hanging with red lanterns in Shennongjia Nature Reserve

Figure 8. Internal structure model of the distinctive building in Shennongjia Nature Reserve

5. CONCLUSIONS

In this paper, a portable four-camera stereo photographic measurement system is developed and implemented, which can provide a professional solution for fast 3D data acquisition, processing, reconstruction and visualization. Compared to traditional methods, it is a non-contact measurement method and generates 3D models containing detailed descriptions of both its appearance and its internal structure, including architectural components.

The system is characterized by: (1) it is portable and can be carried and operated by one person conveniently; (2) multi-source data can be integrated to fully automatically generate 3D models; (3) high-efficient modelling, once images of the scene are captured by the four cameras, they will be transmitted to computers to generate models through the Internet in real-time;

(4) indoor-outdoor seamless modelling, it reconstructs 3D models containing detailed descriptions of both its appearance and its internal structure.

The system is applied to the indoor-outdoor seamless modelling of distinctive architecture existing in two typical cultural tourism zones, that is, Tibetan and Qiang ethnic minority villages in Sichuan Jiuzhaigou Scenic Area and Tujia ethnic minority villages in Hubei Shennongjia Nature Reserve, providing a new method and platform for protection of minority cultural characteristics, 3D reconstruction and cultural tourism.

ACKNOWLEDGEMENTS

REFERENCES

Furukawa, Y. and J. Ponce, 2010. Accurate, dense, and robust multiview stereopsis. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 32(8): p. 1362-1376.

Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. International journal of computer vision, 60(2), 91-110.

Koutsoudis, Anestis, et al, 2014. Multi-image 3D reconstruction data evaluation. Journal of Cultural Heritage, 15(1), 73-79.

Robertson, D.P. and R. Cipolla, , 2009. Structure from Motion. *Practical Image Processing and Computer Vision*. John Wiley, Hoboken, NJ, USA,p. 49.

Snavely, N., 2008. Bundler: Structure from motion for unordered image collections.

Wu, C., 2011. 9. VisualSFM: A visual structure from motion system. http://homes. cs. washington. edu/~ ccwu/vsfm.