

Research Article

Modeling Virtual Reality Environment Based on SFM Method[#]

Jiwu Wang*, Chenyang Li, Shijie Yi, Zixin Li

*School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing, China***ARTICLE INFO***Article History*

Received 25 October 2018

Accepted 09 December 2018

*Keywords*Virtual reality
SFM method
modeling
unity**ABSTRACT**

Virtual Reality (VR) technology is widely used in digital cities, industrial simulation, training, etc. where the environment modeling is a necessary component. Comparing with the difficulty to build the three dimension (3D) environment with the conventional methods, the Structure from Motion (SFM) method is proposed in this paper. The modeling accuracy is studied by comparing with the real dimensions. The results show the SFM method can give a high precision reconstructed 3D model in a short time.

© 2019 The Authors. Published by Atlantis Press SARL.

This is an open access article distributed under the CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>).

1. INTRODUCTION

Virtual Reality (VR) is an artificial media space built by computer. It makes people enter a virtual environment by interaction devices of multimedia sensors. Using external devices, such as helmets and data gloves, users see an interactive and dynamic 3D views formed by multi-source information. VR has been applied in many fields such as simulated military training, digital museum, gaming, etc. [1].

Modeling the virtual world, that is, the construction of VR environment, is a necessary step. Because of the complexity of the real world environments, it is tedious and difficult to model the working environment in any real world. So many people have to simplify the model process by construct the environments with only necessary components for conventional methods. In this paper, the Structure from Motion (SFM) [2] method is studied to build the 3D environments. The results show it can provide more accurate and vivid 3D model and reduce the modeling time.

Till now, many modeling methods are developed to build the 3D environments for VR, such as:

- (1) Using VB, C++, OpenGL graphics library and other tools to build. These methods generally are inconvenient and require much time to study.
- (2) Using VR software, such as WorldToolKit (WTK) and MultiGen Creator, to model complex 3D graphics. Dang [3] completed the creation of virtual campus environment using MultiGen Creator. By this method, the actual campus scene is basically restored. But the model was relatively simple and not realistic enough so the user cannot feel sufficient immersion. And it is very difficult and time-consuming too.

- (3) Using modeling software and Virtual Reality Modeling Language (VRML) to build, such as 3Dmax and SolidWorks, etc. Yuan et al. [4] used SolidWorks and VRML to modeling the VR environment. In this situation, the modeling software could modeling the object accurately, but it is also not suitable for modeling the real environment as not realistic enough.

2. APPLICATION OF SFM METHOD IN VR MODELING

2.1. The SFM Method

In the SFM method, the camera parameters and three-dimensional environment information are recovered with geometry methods. Taking a bear model as an example, its main process is shown in Figure 1.

Structure from motion method is mainly divided into four steps in Figure 1:

- (1) Input and align photos: Input the photos of the captured model or scene into the SFM program to complete the alignment of the images. In the process of alignment, the feature points whose shape or texture information is more prominent in the photos are extracted first. And then the feature points between the photos are matched. After matching is completed, the positional relationship between the cameras is calculated by different feature points in the photo.
- (2) Generate sparse point cloud: The three-dimensional coordinates of each feature point can be solved by the basic principle of stereo vision. Then, according to the texture information of each feature point, a point cloud with color information is generated.

^{*}Corresponding author. Email: jwwang@bjtu.edu.cn; 17121244@bjtu.edu.cn[#]This work was supported by Project KMGY318002531.

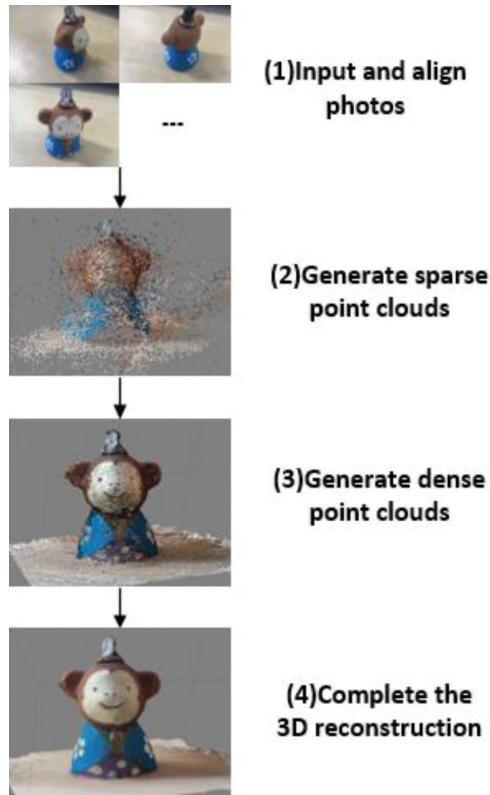


Figure 1 | The process of SFM. (1) Input and align photos. (2) Generate sparse point clouds. (3) Generate dense point clouds. (4) Complete the 3D reconstruction.

- (3) Generate dense point cloud: Using the PMVS tool to get the three-dimensional coordinates of points which are around the feature points and in high photo-consistency, thereby the dense point cloud is generated.
- (4) Complete the 3D reconstruction: After generating dense point cloud, the outliers in the generated dense point cloud need to be removed firstly, the culling operation ensures that the final model has no defects. Then the three adjacent points are used to form a facet, and those facets are connected to form the final 3D model.

One advantage of the SFM method is that the modeling steps are automatic. Another is that it can modeling large-scale scenes, which is very suitable for modeling of natural terrain and urban landscape.

But as SFM method restore the shape and texture information of scene from the feature points in the photos, the method may lead to a bad result in the following cases. One is the environment is low texture, such as a blank wall. There are not enough feature points to extract, so the scene cannot be restored. Another is the problem of shade, some structure cannot be extracted feature points when they were sheltered by other object or other part of their own. Users need to take photos from as many angles as possible in order to avoid that problem.

2.2. Accuracy Analysis of Model Reconstructed by SFM Method

Due to the high requirements on the dimensional accuracy of the model in VR, the modeling accuracy of SFM method is tested with an experiment in this section.

Figure 2 shows a stone table model reconstructed in this experiment. In order to facilitate the dimension measurement in later step, several markers with different number were pasted in the scene.

By comparing the distance in the model and the actual scene, the modeling accuracy of SFM method is analyzed. **Table 1** gives the error analysis data.

According to the data in **Table 1**, error analysis is carried out furtherly in horizontal and vertical directions of the model. Error is mainly classified into Mean Error (ME), Mean Absolute Error (MEA), Standard Deviation of Error (SDE) and Root Mean Square Error (RMSE) in **Table 2**.



Figure 2 | A stone table model.

Table 1 | Error of every distance

Distance between two target	Place that distance exist	Distance in actual scene (mm)	Distance in the model (mm)	Error (mm)	Relative error (%)
Target 1_target 2	Desktop	85	84.874	-0.126	-0.15
Target 5_target 6	Ground	85	85.378	0.378	0.44
Target 13_target 14	Vertical	85.2	84.945	-0.255	-0.30
Target 15_target 16	Vertical	85	84.774	-0.226	-0.27
Target 19_target 20	Vertical	85.2	84.853	-0.347	-0.41
Target 1_target 3	Desktop	403	406.888	3.888	0.96
Target 1_target 4	Desktop	402	405.723	3.723	0.93
Target 1_target 7	Desktop	677	675.643	-1.357	-0.20
Target 1_target 8	Desktop	677	675.623	-1.377	-0.20
Target 2_target 3	Desktop	470.3	468.954	-1.346	-0.29
Target 2_target 4	Desktop	479.5	477.775	-1.725	-0.36
Target 2_target 7	Desktop	651	649.118	-1.882	-0.29
Target 2_target 8	Desktop	661	659.561	-1.439	-0.22
Target 3_target 4	Desktop	85	84.937	-0.063	-0.07
Target 3_target 7	Desktop	624	622.34	-1.66	-0.27
Target 3_target 8	Desktop	572	571.684	-0.316	-0.06
Target 4_target 7	Desktop	705	703.709	-1.291	-0.18
Target 4_target 8	Desktop	652	655.209	3.209	0.49
Target 5_target 11	Ground	668.2	666.842	-1.358	-0.20
Target 5_target 12	Ground	720	718.115	-1.885	-0.26
Target 6_target 11	Ground	731	729.503	-1.497	-0.20
Target 6_target 12	Ground	777	774.876	-2.124	-0.27
Target 7_target 8	Desktop	85	84.85	-0.15	-0.18
Target 9_target 10	Ground	85.1	84.587	-0.513	-0.60
Target 9_target 11	Ground	618.5	617.934	-0.566	-0.09
Target 9_target 12	Ground	554	553.389	-0.611	-0.11
Target 10_target 11	Ground	563	561.864	-1.136	-0.20
Target 10_target 12	Ground	492	491.677	-0.323	-0.07
Target 11_target 12	Ground	85.1	84.938	-0.162	-0.19

Table 2 | Model error analysis

	Horizontal	Vertical	Totality
ME (mm)	-0.498	-0.2865	-0.432
MEA (mm)	1.400	0.287	1.204
SDE (mm)	1.665	0.081	1.525
RMSE (mm)	1.738	0.293	1.585

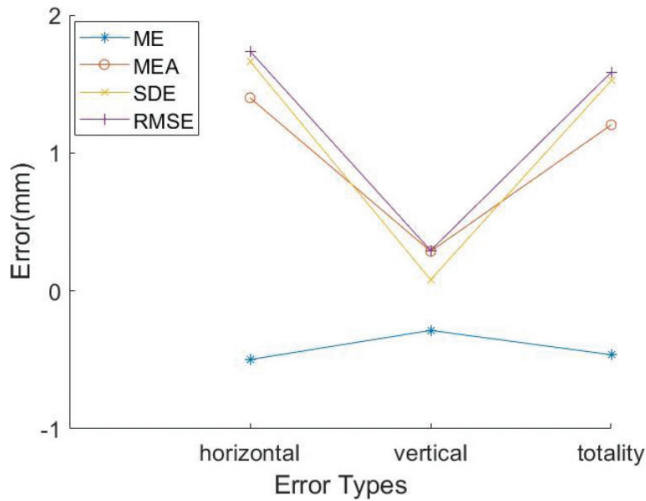
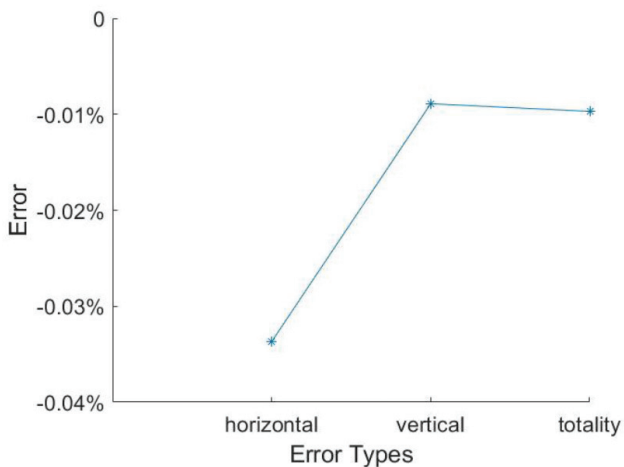
**Figure 3** | Error in all directions.**Figure 4** | Average relative error in all directions.

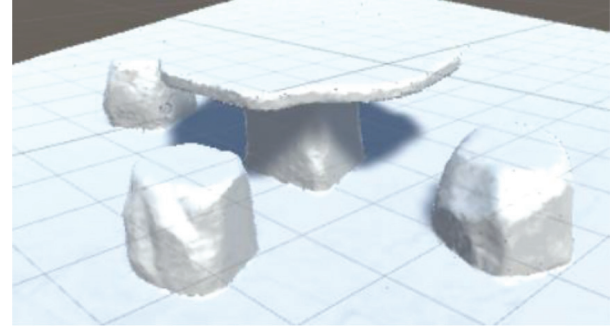
Figure 3 shows the trend of the error in each direction, and Figure 4 shows the trend of the average relative error in each direction.

It can be seen from Figure 2, Tables 1 and 2 that the model reconstructed is basically the same as the actual scene. From Figures 4 and 5, it can be concluded that the error value of the horizontal and vertical directions is low. But the error value of the horizontal direction is larger as the distance between makers in the horizontal direction is larger. After comprehensive analysis, the final average error of the model is <1 mm, the maximum error is <4 mm, and the average relative error is <0.1%. Moreover, the modeling time is all within 2 h.

2.3. Application of SFM Method in VR Modeling

After the modeling is completed, the material information can be exported through a 3D format file such as OBJ, and the texture file of the model is also exported as a picture. The exported model and texture information are respectively shown in Figures 5 and 6.

The files obtained in the above steps can be imported into the Unity3D software successively to complete the subsequent edit of the VR. The imported model in Unity3D is shown in Figure 7.

**Figure 5** | The material information of model.**Figure 6** | The texture information of model.**Figure 7** | The model in Unity3D.

3. CONCLUSION

Using SFM method to modeling the VR environment based on real environment is discussed in this paper. Compared with conventional modeling methods, the advantages of SFM method are shown as follows:

- (1) Reduce the difficulty of modeling by less operations. Using the SFM method, the modeling only needs to take a multi-angle photos of the real environment in the early stage, and then input photos into the program, in which the modeling can be operated step-by-step. Compared with the conventional methods, the operation is very simple and requires little on modeling skill in SFM method.
- (2) Reduce the modeling cycle with less modeling steps. The SFM method needs less operation steps than the conventional method, and the time period is also shorter, even within 2 h, which can be even faster on a computer with a high-performance GPU.
- (3) It is more suitable for the VR modeling. Compared with conventional method modeling, the texture information modeled by the SFM method extracts from a photo rather than a manual

selection of tones, so a higher degree of realism can be obtained when the texture is finally mapped onto the model. And as discussed in [Section 2.2](#), the error of model reconstructed by SFM method is low enough to meet requirements of VR.

On the other hand, there are also two cases to be attention: one is that the SFM method should not be used to modeling the scenes in low texture; another is that users should take photos of the scenes in as many angles as possible.

REFERENCES

- [1] X. Shuo, M. Kun, L. Shuqin, D. Meng, W. Lijun, Research on key technologies of online experiential application based on VR/AR, *Intell. Comput. Appl.* 7 (2017), 28–31.
- [2] L. Shapiro, G. Stockman, *Computer vision*, Prentice Hall, Upper Saddle River, New Jersey, 2001.
- [3] B. Dang, *Research on virtual scene generation in virtual reality*, Beijing Jiaotong University, Beijing, 2010.
- [4] F.W. Yuan, B.W. Li, B. He, Accurate modeling of VR based on SolidWorks-VRML, *Mech. Electr. Eng. Magazine* 27 (2007), 103–105.

Authors Introduction

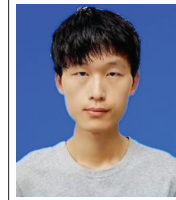
Dr. Jiwu Wang



He is an Associate Professor at School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, China.

He obtained his PhD degree at Tsinghua University, China in 1999. His main interests are computer vision, robotic control and fault diagnosis.

Mr. Shijie Yi



He is a master student at School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing China. His main interest is computer vision for robotic.

Mr. Chenyang Li



He is a master student at School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing China. His main interests are 3D reconstruction and object detection.

Ms. Zixin Li



She obtained her master degree at School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, China in 2019. Currently, she works at AECC Beijing Hangke Control System Science and Technology Co., Ltd. Her main interest is 3D reconstruction.