The 13th International Conference on
Computer Science & Education (ICCSE 2018)
August 8-11, 2018. Colombo, Sri Lanka

Research on 3D Reconstruction Based on Multiple Views

Xu Chen
Department of Software Engineering
Xiamen University
Xiamen, China
1026046241@qq.com

Qingfeng Wu*
Department of Software Engineering
Xiamen University
Xiamen, China
qfwu@xmu.edu.cn

Shengzhe Wang
Department of Software Engineering
Xiamen University
Xiamen, China
1005709383@qq.com

***Abstract*—The acquisition of 3D models is a fundamental research in the field of computer graphics and computer vision. However, it is very complicated and expensive to build 3D models manually by using modeling software, such as 3DMAX and Maya. Therefore, how to get 3D models directly and quickly from the real world is a hot research topic in this field. Multi-view based 3D reconstruction techniques include feature point extraction and matching, camera calibration, sparse point cloud reconstruction, dense point cloud reconstruction, Poisson surface reconstruction, texture mapping and many other techniques. This paper will focus on the main algorithms of 3D reconstruction algorithms - Structure from Motion, Multi-View Stereo and Poisson Surface Reconstruction. Based on the above research, this paper designed a 3D reconstruction system including the above algorithms. The system takes a variety of views of an object as input, and finally outputs the 3D model of the object. Finally, this paper verified the effectiveness and feasibility of the system through experimental analysis.**

Keywords—3D Reconstruction, Structure from Motion, Multi-View Stereo, Poisson Surface Reconstruction

I. INTRODUCTION

The traditional 3D modeling technology is the technology of building 3D model of the object by using the modeling tools such as 3DMAX and Maya and using the knowledge of computer graphics. Because of its high accuracy and complete texture information, it is widely used in 3D animation and other fields.

But in the modeling process, the texture of the model must rely on manual paste, which makes the modeling cycle and workload increase greatly. Especially in the establishment of a real world, such as schools, parks, and so on, the traditional modeling technology can't truly restore the hidden details in the scene, which makes the traditional modeling technology be impacted. Therefore, how to get 3D models directly and quickly from the real world is a hot topic in this field.

At present, the acquisition of 3D structure of real objects is a digital storage and recording technology, which has extensive application requirements in the scientific and engineering fields, such as object modeling, scene modeling, realistic rendering, robot navigation, target recognition and 3D measurement, as well as other cultural domains, such as archaeology, advertising, entertainment and so on. To sum up, the research of 3D reconstruction technology based on multi-view has great theoretical research value and practical application significance.

The rest of this paper is organized as follows. The related work of 3D reconstruction is briefly reviewed in Section 2. In Section 3, 3D reconstruction algorithms are proposed. In Section 4, simulation results and performance comparisons are presented. Finally, the conclusions of our analysis are presented in Section 5.

II. RELATED WORK

In the field of computer vision, the problem of restoring the motion parameters of the camera and the three-dimensional geometry of the space object from multiple images is called the 3D reconstruction based on multi-view.

At present, there are a lot of 3D reconstruction algorithms. Tomasi and Kanade [1] assume that the camera is the orthogonal projection model, and use the affine decomposition method to solve the 3D structure and camera motion simultaneously. Debevoc [2] designed the famous building reconfiguration system Facade. The system requires the rough geometric model of the building and the motion parameters of the camera first, then the model is projected back to the image and compared with the actual image, and the exact three-dimensional structure of the building is finally calculated by reducing the back projection error. Pollefeys [3] have successfully applied self-calibration and hierarchical reconstruction to archaeology, cultural relic protection and other fields, and achieved good results. HY.Shum [4] proposed a human-computer interactive reconstruction system that could restore a three-dimensional structure from a set of panoramic mosaic, or the scene as a set of sprite divided by depth. Faugeras [5] uses hierarchical reconstruction, self-calibration and other methods to reconstruct buildings from image sequences.

III. MAIN ALGORITHMS

The 3D reconstruction technology based on multi-view is composed of techniques, such as feature point extraction and matching, camera calibration, sparse point cloud reconstruction, dense point cloud reconstruction, Poisson surface reconstruction, texture mapping and so on [6]. The input is a set of pictures and the output is a 3D model. The main algorithms include Structure from Motion, Multi-View Stereo and Poisson Surface Reconstruction. The following three algorithms will be introduced.

A. Structure from Motion

Structure from Motion (SfM) originated from photogrammetry. It is one of the crowning achievements of photogrammetry and computer vision [7].

The implementation of SfM algorithm is mainly composed of four steps: feature extraction, feature matching, path generation and incremental reconstruction [8]. The incremental SfM algorithm process is shown in Fig. 1.

978-1-5386-5495-8/18/$31.00 ©2018 IEEE

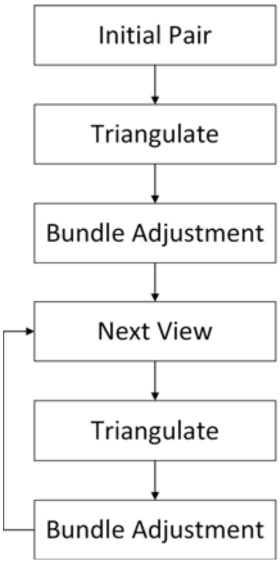

Fig. 1. The incremental SfM algorithm process

1) Feature matching: Searching for interest points that can be distinguished from other images in every picture. Around these points, it is extracted and stored in a feature descriptor. In the picture, variables require the invariance of feature descriptors, which are related to image scaling, rotation, noise, exposure, contrast and so on.

2) Feature extraction: Using the nearest neighbor method to find the smallest Euclidean distance in a high dimensional space, so as to find a corresponding descriptor for each descriptor in another picture.

3) Path generation: The 22 pairing of feature descriptors generates many paths. These paths can be combined into multiple views to form characteristic paths.

4) Incremental reconstruction: Incremental reconstruction will be conducted under the guidance of initialized camera pairs. It's very important to find a camera pair to avoid configuration degradation. In general, a perfect camera pair has many pairs of matches and a lot of parallax. If the distance between two cameras is too small, triangulation will become unstable and lead to poor 3D point set. At the same time, if many matches correspond to the plane area corresponding to the scene, the focal length and depth of field of the camera will be no difference.

B. Multi-View Stereo

The images obtained through SfM are overlap. In order to reduce the amount of data and improve the efficiency of dense matching, the Cluster Multi-View Stereo algorithm is needed to classify the images. Then, the final dense matching is completed by the three main steps of matching, expansion and filtering by the Patch-based Multi-View Stereo and the dense point cloud is generated.

1) Cluster-based Multi-View Stereo: Cluster-based Multi-View Stereo algorithm contains the following four steps, see Fig. 2.

a) SfM filtering: Searching the visual information of a SFM feature point in the local neighborhood, and then taking the location information as the average value of each neighborhood location. This can effectively reduce the number of input point sets. The step is repeated, and the final output point set is made up.

b) Image selection: Delete the image which is not satisfied with the constraint according to the coverage scope mentioned above. It is noted that the order of search is gradual search according to image resolution from low to high, so that low resolution images can be first eliminated.

c) Clustering classification: By standard segmentation algorithm, the image size is constrained, and the coverage constraint is not segmented. The clusters that do not meet the size of the image are divided into smaller clusters.

d) Augmented image: For each SfM feature point that is not added, each SfM feature point corresponds to a unique efficiency value, and the final selection of the maximum efficiency value corresponding to the point is added to the cluster.

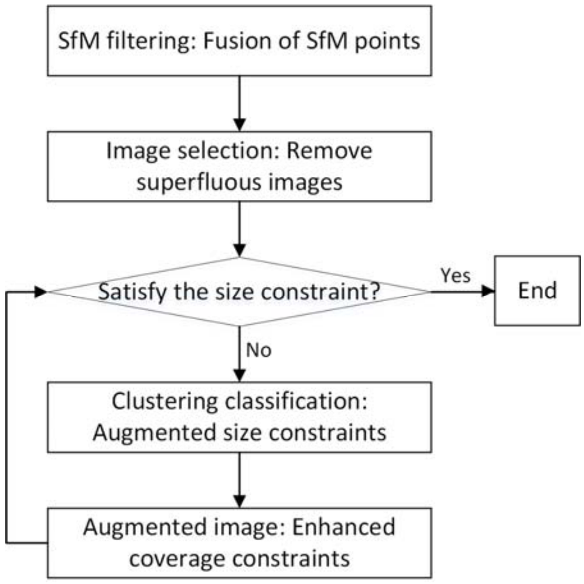

Fig. 2. Clustering algorithm implementation process

2) Patch-based Multi-View Stereo: Patch-based Multi-View Stereo is a good quasi-dense 3D reconstruction method based on multi-view [9]. The algorithm is composed of matching, expansion and filtering process, starting from the sparse point set composed of key points, spreading the key points, and filtering out the error matching point by visual constraint.

a) Initial feature matching: Firstly, extract the feature points of each image. For each feature point in each image, the matching point pair is formed by finding the same type of feature points in the other images by allowing the polar line constraints of two pixel errors. Then we use these matching points to generate a series of 3D spatial points using triangulation method, and finally try to generate patches in turn.

b) Patch expansion: The purpose of patch expansion is to ensure that each image block corresponds to at least one patch. Through the patch generated above, the new patch is generated repeatedly.

c) Patch filtering: During the reconstruction process, we may generate some patches with serious error, so we need to filter them to ensure the accuracy of the patches.

C. Poisson Surface Reconstruction

Poisson Surface Reconstruction (PSR) realizes the process of generating polygonal mesh from dense point clouds [10]. PSR can produce smooth surface and is robust to noise.

The 3D surface reconstruction technique based on Poisson equation is used to study the algorithm for the reconstruction of triangular surface model with the geometric surface information of the object on the basis of the three dimensional point cloud model which has the information of the normal vector. The steps of the whole algorithm include the preprocessing of the information of the input point cloud with the normal vector information, the discretization of the global problem, the solution of the discrete sub data, the extraction of the equivalent surface after the Poisson problem, and the later optimization [11]. For the 3D surface reconstruction algorithm, the Poisson surface reconstruction algorithm combines the advantages of the global and the local method. By using the implicit fitting method, the implicit equation represented by the surface information described by the point cloud model is obtained by solving the Poisson equation. By extracting the equal value surface of the equation, the surface model which has the rich three-dimensional geometric entity information is obtained. The model reconstructed by this method has the closed characteristics of watertight, and has good geometric surface features and detail features.

The steps of the whole algorithm include the preprocessing of the information of the input point cloud with the normal vector information, the discretization of the global problem, the solution of the discrete sub data, the extraction of the equivalent surface after the Poisson problem, and the later optimization.

The process of surface reconstruction is divided into five steps [12].

1) Define octree: The octree is used to store the point set, and the octree is defined according to the location of the sampling point set, and then the octree is subdivided into the leaf nodes with depth of D.

2) Set function space: Set space function F for each node of octree, the linearity of all node function F and vector field V, base function F use box filtering n-dimensional convolution.

3) Creating vector field: Under the condition of uniform sampling, the block is assumed to be constant, and the gradient of indicator function is approximated by vector field V. Three spline interpolations are used.

4) Solving Poisson's equation: The solution of the equation is solved by Laplasse matrix iteration.

5) Extracting the contour surface: In order to get the reconstructed surface, we need to select the threshold to obtain the equivalent surface. First estimate the location of the sampling point, then use its average value to extract the equivalent surface, and then use the moving cube algorithm to get the equivalent surface.

IV. Experiment

We developed and tested our multi-view 3D reconstruction system under Windows. The system takes a variety of views of an object as input, and finally outputs the 3D model of the object. In the following, we show results on a few datasets we acquired over time.

Kermit: The first data set is called Kermit. It contains 9 images of a plush toy and some other objects. This data set is obtained from the network. Fig. 3 shows the 9 input images, and Fig. 4 shows the sparse point cloud, dense point cloud, polygonal mesh and final 3D model generated during the operation of the system.

Fig. 3. The data set of Kermit

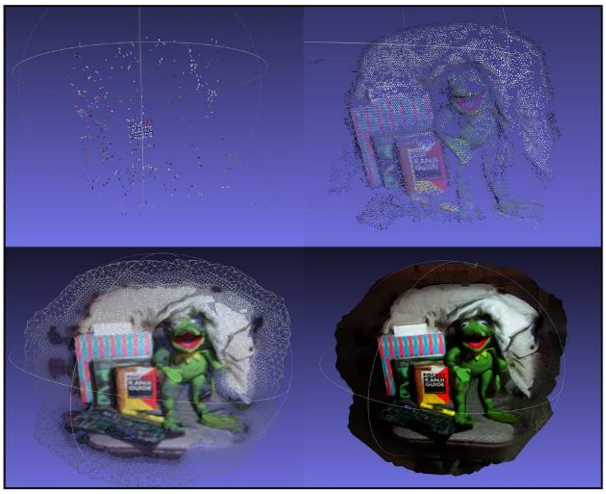

Fig. 4. The sparse point cloud (upper left), dense point cloud (upper right), polygonal mesh (lower left) and 3D model (lower right)

Der Hass [7]: The next data set is called Der Hass and it is also obtained from the network. It contains 79 images of a massive stone sculpture. This is a relatively compact data set with uniform scale as the images have the same resolution and are evenly spaced around the object. Fig. 5 shows the angle of the images sampling of the object. Fig. 6 and Fig. 7 show a dense point cloud and a 3D reconstruction model of the Der Hass data set.

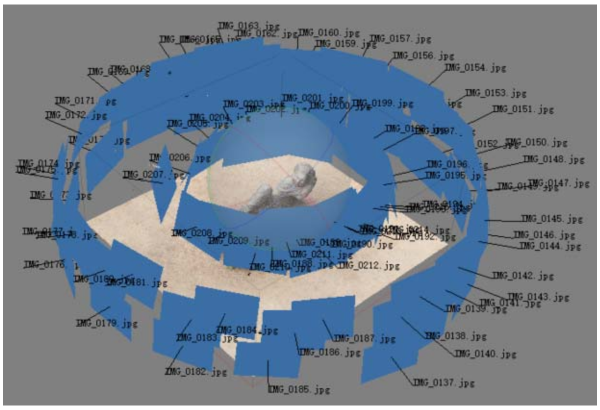

Fig. 5. The angle of the images sampling of Der Hass

Fig. 6. Dense point cloud (left) and 3D reconstruction model (right) of Der Hass

Fig. 7. Dense point cloud (left) and 3D reconstruction model (right) of Der Hass

Street View: Next, we reconstruct the street view from 182 images captured by UAV. We don't capture it in line with a specific rule, which made the dense points cloud appear some holes, see Fig 8.

Fig. 8. Dense point cloud of street view

QunXian Buildings: We conclude our demonstration with the QunXian Buildings data set in Fig. 9. The 268 input images depict old historic buildings. It generates more than 500 thousand valid points and more than 1 million faces.

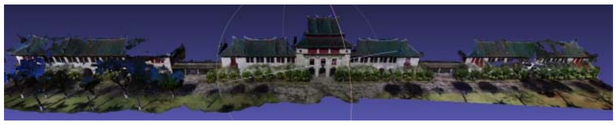

Fig. 9. Dense point cloud of QunXian Buildings

The results show that our system is multi-functional and can operate on a broad range of data sets. It's not only the quantity of Images, but the angle and quality of capturing determines the quality of 3D models. It means that a densely sampled spiral around compact objects with a large overlap between the photos leads to the best results. A sparse sampling may lead to disconnected components in SfM or holes in the MVS reconstruction. We can conclude the rules of capturing pictures from experiments, see Fig. 10.

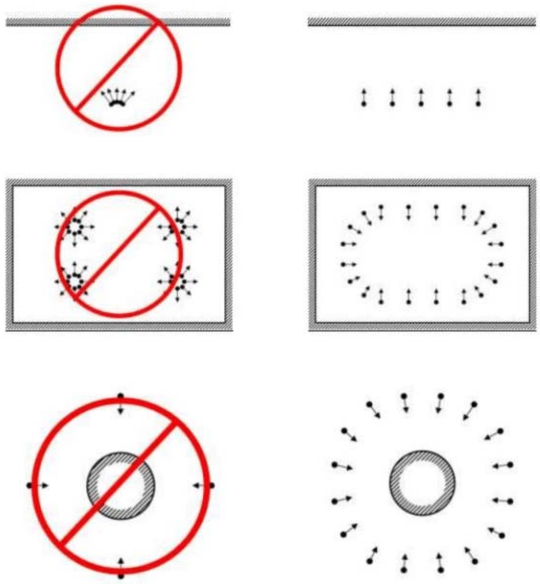

Fig. 10. The rules of capturing pictures

V. Conclusion and Expectation

Multi-view based 3D reconstruction technology has extensive research and application background in computer graphics, computer vision, virtual reality and other fields. This paper focuses on the main algorithms of 3D reconstruction technology based on multi-view. Based on the above algorithms, we design a 3D reconstruction system. It takes a variety of views of an object as input, and finally outputs the 3D model of the object. We verified the effectiveness and feasibility of the system through experimental analysis. Experiments show that the system is multi-functional and can operate on a broad range of data sets. However, a practical limitation in the presented system is the memory consumption in some parts of the pipeline. The more pictures input, the more memory need to occupy. We plan to accelerate the part of feature matching by GPU to make our system more efficient and perfect.

Acknowledgment

This work was supported by NSFC (No. 61402387, No. 61402390); the Key Program of Science and Technology of Fujian Province of China (No. 2014H0044); Science and Technology Guiding Project of Fujian Province of China (No. 2015H0037, No. 2016H0035); Enterprise Technology

Innovation Project of Fujian Province; the Education and Research Project of Middle and Young Teacher of Fujian Province of China (No. JA15018); the Overseas Study Scholarship of Fujian Province; Science and Technology Project of Xiamen, China (No. 3502Z20153026); Research and Application of Artificial Intelligence Technology Based on Clinical Big Data of Tongue Diagnosis (No. 2017YFC1703303).

REFERENCES

[1] Tomasi C, Kanade T. Kanade, T.: Shape and motion from image streams under orthography: A factorization method. Int. J. Comput. Vis.9(2), 137-154[J]. International Journal of Computer Vision, 1992, 9(2):137-154.

[2] Debevec P E. Modeling and rendering architecture from photographs[M]. University of California, Berkeley, 1996.

[3] Pollefeys M. Self-calibration and metric 3d reconstruction from uncalibrated image sequences[J]. Thesis K.u.leuven Departement Esat Afdeling Psi.phd.thesis, 1999.

[4] Shum H Y, Szeliski R, Baker S, et al. Interactive 3D Modeling from Multiple Images using Scene Regularities[C]// IEEE Workshop on Applications of Computer Vision. IEEE Computer Society, 1998:234.

[5] Faugeras O, Laveau S, Robert L. 3-D Reconstruction of Urban Scenes from Sequences of Images[M]// Automatic Extraction of Man-Made Objects from Aerial and Space Images. Birkhäuser Basel, 1995:145-168.

[6] Simon Fuhrmann, Fabian Langguth, Nils Moehrle. MVE—An image-based reconstruction environment[J]. Computers & Graphics, 2015, 53(PA):44-53.

[7] Smith M W, Carrivick J L, Quincey D J. Structure from Motion Photogrammetry in Physical Geography[J]. Progress in Physical Geography, 2016, 40(2).

[8] Fuhrmann S, Langguth F, Goesele M. MVE - A Multi-View Reconstruction Environment[J]. 2014.

[9] Li M, Zheng D, Zhang Y. An Improved Patch-Based Multi-View Stereo Algorithm for Large Image Sets[J]. Journal of Computational & Theoretical Nanoscience, 2016.

[10] Kazhdan M, Bolitho M, Hoppe H. Poisson surface reconstruction[C]// Eurographics Symposium on Geometry Processing. 2006:61-70.

[11] Kazhdan M, Hoppe H. Screened poisson surface reconstruction[J]. Acm Transactions on Graphics, 2013, 32(3):1-13.

[12] Hoppe H. Poisson surface reconstruction and its applications[C]// ACM Symposium on Solid and Physical Modeling. ACM, 2008:10-10