

# 3D MEASUREMENT AND MODELING BASED ON STEREO-CAMERA

ZHENG Shunyi<sup>1</sup>, WANG Ruirui<sup>1</sup>, CHEN Changjun<sup>2</sup>, ZHANG Zuxun<sup>1</sup>

<sup>1</sup> School of Remote Sensing and Information Engineering, Wuhan University, 430079, Luoyu Road 129, Wuhan, Hubei, P.R.China

<sup>2</sup> School of Geodesy and Geomatics, Wuhan University, 430079, Luoyu Road 129, Wuhan, Hubei, P.R.China

Commission V, WG V/1

**KEY WORDS:** Cultural Heritage, Camera, Modelling, Measurement, Registration, Reconstruction, Three-dimensional

## ABSTRACT:

An automatic 3D measurement and modelling method based on stereo-camera was proposed in this paper. Firstly a hardware system for data capturing, named stereo-camera system, was developed. This system included three cameras, which were fixed by joint poles between them. Cameras' intrinsic parameters and relatively extrinsic parameters between them could be obtained by calibration. With the help of stereo-cameras three images of the object can be captured at the same time and then processed with digital photogrammetric method to produce local point cloud. Meanwhile texture data could be picked up from the middle image and the textured local 3D surface model can be constructed. All local models were produced by the same process. Then a whole adjustment optimization method was developed to register all the local models into a uniform coordinate system with the corresponding point information between adjacent local models, and finally the entire 3D model could be constructed. This method has the same function as laser scanner. However, compared with laser scanner, this method is low-cost. In addition, texture can be extracted from original images and textured model can be reconstructed easily because it is image-based method. Instead of two cameras, this system used three cameras, so high reliability can be achieved. Its main shortcoming is that sometimes it fails for objects lacking texture. Experiment results prove above conclusions.

## 1. INTRODUCTION

Building models of spatial objects based on 2D images is the active topic of research in the fields of photogrammetry and computer vision. By now photogrammetry technique has broad application in many fields, especially for map production, which is one of the most successful examples of 3D measurement and modelling based on images. Recently, 3D measurement and modelling based on images which are supplied by digital camera not only has boarder applications but also has successfully exploited many new application fields such as preserving cultural heritage, measuring industrial parts, and 3D measuring and modelling of people's faces and body due to the lower cost. The process of existing modelling methods is very complicated, and the models which are constructed by these methods have lower accuracy, or have a high cost when the accuracy satisfies the requirement such as laser scanner method. So how to construct models with a higher accuracy fleetly and automatically is a problem to be solved in time.

This paper proposes a new automatic 3D-scene modelling method based on stereo-camera system. Firstly, a hardware system for data capturing was designed, named stereo-camera system, which included three cameras. With the help of stereo-camera system three images of the object could be gotten simultaneously, which create a 3D local model with high accuracy. This system has the same function as laser scanner but with a much lower cost. After obtaining image data, the local 3D surface model was firstly constructed, and then all the local models were registered to a whole surface model on the basis of the correspondent points between adjacent local models by virtue of a whole adjustment optimization method which

improves the accuracy of registration. The 3D surface modelling method in study has been proved to be effective and practical by the experimental data and results. This method has the distinct advantage of low cost and high quality.

## 2. HARDWARE SYSTEM FOR DATA CAPTURING

In order to make capturing image data simple and automatic, this paper designed a hardware system for data capturing, which was called stereo-camera system (Figure 1). This system included three cameras which were fixed by joint poles between them. By calibration the intrinsic parameters of every camera and relatively extrinsic parameters of left and right cameras relative to middle camera could be obtained. With the help of this system three images of every local scene could be captured simultaneously, which named local image set. Adjacent local image set must have certain overlap area in order to make sure accurate registration. All the local image sets of the whole scene were arranged in row order. This system has low cost because that it mainly composes of three cameras. Compared with stereo photopair formed by two images captured by two cameras, two stereo photopairs could be formed simultaneously from the local image set, which supply high reliability to the modelling. Apparently, the stereo-camera system has the same function as laser scanner which can capture range data of high quality. Considering the 3D laser point clouds are distributing densely and equably because of the equality of laser point emission, it lacks the sufficient expression of scene features. Additionally, the laser scanner has a high cost at present. As a result, compared with laser scanner, the stereo-camera system would have a board application due to low cost and high quality.

### 3. THE CONSTRUCTION OF 3D SURFACE MODEL

After obtaining image data, the next step was to create the 3D surface model. In this paper, the local surface model was firstly constructed on the basis of local image sets, then all the local models were registered into a uniform coordinate system to form a whole model with the corresponding point information in the overlapped areas, lastly the entire 3D surface model could be constructed. The flowchart of this method is depicted as figure 2.

#### 3.1 The construction of local scene model

Reconstruction of 3D photorealistic scene model includes two aspects: the first step is to get accurate geometry information of scene, and the second step is to capture texture information. In order to get geometry data and texture data digital photogrammetric method and projection method were respectively adopted which are explained in detail as follows.

In the process of digital photogrammetric measurement the direction of two conjugate rays in space can be gotten based on the same image points on stereo photopair. These two conjugate rays must have an intersection point in space which just is the spatial position of the corresponding ground point. Two stereo photopairs could be formed simultaneously on the basis of every local image set which were processed after image distortion rectification. By inputting the respective intrinsic parameters of three cameras, relatively extrinsic parameters between them and image coordinate values of homologous points on three images, the error equation and normal equation could be constructed with respect to the collinearity equation (Equation 1), so the 3D point clouds of local scene could be obtained by least-square method. But photorealistic scene surface model is not just the set of feature points in the scene, which also needs texture mapping with texture information of surfaces.



Figure 1. The stereo-camera system

$$\begin{aligned} x - x_0 &= -f \frac{a_1(X - X_s) + b_1(Y - Y_s) + c_1(Z - Z_s)}{a_3(X - X_s) + b_3(Y - Y_s) + c_3(Z - Z_s)} \\ y - y_0 &= -f \frac{b_1(X - X_s) + b_2(Y - Y_s) + c_2(Z - Z_s)}{a_3(X - X_s) + b_3(Y - Y_s) + c_3(Z - Z_s)} \end{aligned} \quad (1)$$

where  $f$  is focus length of camera

$x, y$  is image coordinates

$X_s, Y_s, Z_s$  are coordinates of projection center

$X, Y, Z$  are object coordinates in ground coordinate system

$a_i, b_i, c_i$  ( $i=1, 2, 3$ ) are direction cosines formed by the three angles ( $\varphi, \omega, \kappa$ ), which also are named orthogonal rotation matrix.

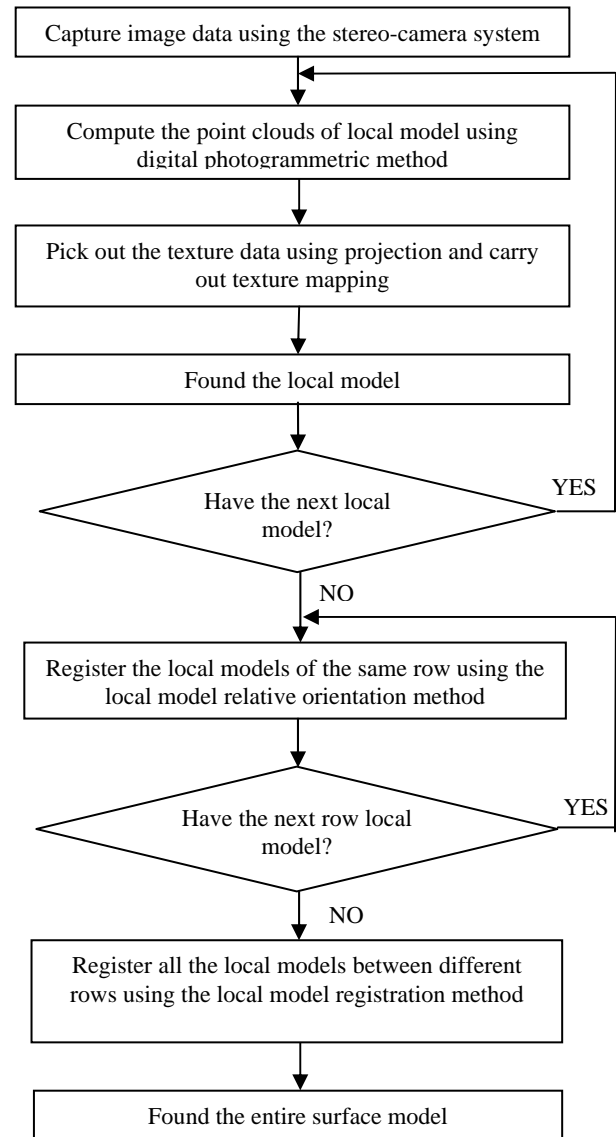


Figure 2. The modeling flowchart based on the stereo-camera system

The next step was to capture the texture data of every relevant plane in the range of projecting plane by projecting every geometrical plane to the middle image in the local image set, and then carried out texture mapping with the help of software algorithm. By now, the whole surface model was completely constructed. In addition, sometimes the texture of certain scenes which was automatically picked up could have a bad effect because of occlusion. The deficiency could be remedied with the help of the texture of right and left images or additional photos which were taken in addition.

After every local image set were processed by the method which is introduced above, all the local scene models could be obtained.

### 3.2 The registration of local models

The registration of local models could be completed by the spatial similar transformation which is without scaled factor. The mainly idea of this method is explained as follows: firstly a coordinate system was assigned as world coordinate system; basing on the correspondence points between adjacent local models, the registration between local models of same rows was implemented, and then all the local models of every row were registered; finally the whole scene model would be constructed.

In the process of registration the local model is regarded as a rigid body which can only be translated and rotated. The transformation can be realized by the spatial similar transformation without scaled factor (Equation 2).

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = R_i \begin{bmatrix} X_{i+1} \\ Y_{i+1} \\ Z_{i+1} \end{bmatrix} + \begin{bmatrix} T_{i,X} \\ T_{i,Y} \\ T_{i,Z} \end{bmatrix} \quad (2)$$

where  $R_i$  is orthogonal rotation matrix formed by the three rotation angles in the process of relative orientation

$[X_i, Y_i, Z_i]^T$  is corresponding feature points' coordinates in local model  $i$

$[X_{i+1}, Y_{i+1}, Z_{i+1}]^T$  is corresponding feature points' coordinates in local model  $i+1$

$[T_{i,X}, T_{i,Y}, T_{i,Z}]^T$  is translation vector

In the process of registration of local models of same rows, the coordinate system of the first local model was regarded as the world coordinate system, the second model which was adjacent to the first one could be registered to the first model by computing the parameters of spatial similar transformation on the basis of correspondent points in the overlapped areas. The rest local model may be processed by analogy method and finally every local model is registered to the previous one, then at last all the local models in the same row could be registered to a uniform coordinate system. In order to get a more accurate result, this paper adopts a whole adjustment optimization method to reduce the accumulate error which was arisen in the process of model registration.

After all the registrations between local models of same rows were completed, the registration between rows could be carried out. Similarly, the coordinate system of the first row local model was regarded as the world coordinate system, the second model which was adjacent to the first one could be registered to the first row model by computing out the parameters of spatial similar transformation on the basis of common feature points in the overlapped areas. Then with the help of the same method every row local model could be registered to the previous one, lastly all the row local models could be registered to a whole one which was the whole scene model. A more accurate scene

model could be gotten after optimizing the scene model by the whole adjustment optimization method.

### 3.3 The whole adjustment optimization of model registration

Considering the accumulate error produced in the registration of models, all the local models were processed by the whole adjustment optimization in order to get a more accurate model.

The mainly idea of the whole adjustment optimization is as follows: firstly based on the correspondence points between all the adjacent local models the 3D similarity transformation formula between every model and the former model is gotten and then processed by linearization, lastly the error equation can be listed as follows:

$$-\begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix}_{ij} = \begin{bmatrix} 1 & 0 & 0 & X & -Z & 0 & -Y \\ 0 & 1 & 0 & Y & 0 & -Z & X \\ 0 & 0 & 1 & Z & X & Y & 0 \end{bmatrix}_{ij} \begin{bmatrix} dT_x \\ dT_y \\ dT_z \\ d\phi \\ d\omega \\ d\kappa \end{bmatrix}_{ij} - \begin{bmatrix} dX \\ dY \\ dZ \end{bmatrix}_{ij} - \begin{bmatrix} l_x \\ l_y \\ l_z \end{bmatrix}_{ij} \quad (3)$$

and

$$\begin{bmatrix} l_x \\ l_y \\ l_z \end{bmatrix}_{ij} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{ij} - R_{ij} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{i+1j} - \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}_{ij} \quad (4)$$

where  $j$  is low of the local model

$i$  is serial number of the local model in the same row image coordinates

$R_{ij}$  is orthogonal rotation matrix formed by three rotation angles  $(\phi, \omega, \kappa)$  in the process of model registration

$[X, Y, Z]_{i+1j}^T$  is coordinates of model  $i+1$  which located in row  $j$

$[X, Y, Z]_{ij}^T$  is coordinates of model  $i+1$  located in row  $j$  in the coordinate system of model  $i$  located in the same row

$[T_x, T_y, T_z]_{ij}^T$  is translation vector

$[V_x, V_y, V_z]_{ij}^T$  is correction coordinates of models' points which assigned as the observation values

$[l_x, l_y, l_z]_{ij}^T$  is constant term of the error equation

$[dX, dY, dZ]_{ij}^T$  is correction coordinates of models' points

$[dT_x, dT_y, dT_z, d\phi, d\omega, d\kappa]_{ij}^T$  is corrections of 6 spatial similar translation parameters between model  $i+1$  located in row  $j$  and the former model located in the same row

The equation 3 and equation 4 are fitted to correspondence points between adjacent models. For control points,  $[dX, dY, dZ]_{ij}^T$  equal to 0 because of control points' error is 0. Based on every normal joint points or control points three error equations can be constructed.

The models' directional parameters and coordinate values could be gotten by computing the above equations, which can improve the accuracy of model registration and make the whole 3D scene model more accurate.

#### 4. EXPERIMENT AND ANALYSIS

In order to prove the method introduced in this paper, experiments were conducted, one of which is to measure an ancient tomb constructed in Chinese Han dynasty. Firstly a stereo-camera system constituted of three digital cameras was designed, and then relatively extrinsic parameters of left and right cameras based on middle camera and the intrinsic parameters of every camera which include focus, principle point and distortion parameters were obtained by calibration. The entire scene was divided into 124 local scenes; in every local scene three images were captured which were regarded as a local image set depicted in Figure 3, so the whole scene included 124 local image sets in all.

The images were processed with digital photogrammetric method to produce local point clouds which were depicted in Figure 4 and texture mapping by projection, lastly the local scene model could be constructed (Figure 5). In order to get the whole scene surface model (Figure 6), a whole adjustment optimization were adopted to register all the local scene models into a uniform coordinate system to form a whole model based on corresponding feature points in the overlapped areas.

The experiment results indicate that the operation of this modelling method is simple and high automatic; meanwhile the final 3D surface model has high quality. However in view of occlusion the texture of certain scenes which was automatically picked up could have a bad effect, manual work must be added to get a better result.



Figure 3: The local image set



Figure 4: The point clouds Figure 5: The local scene model



Figure 6: The entire scene model

#### 5. CONCLUSIONS

This paper addressed the problem of 3D measurement and reconstruction of the 3D scene modelling based on the images captured by non-metric camera, and brings up a solution which includes a hardware system of data capturing and software algorithm. The experiment indicates that the stereo-camera system is of a simple structure and low cost. With the help of this system image data of high quality can be obtained conveniently, and the final 3D surface model has high accuracy. So this method would have a board application prospect. Additionally, the study on texture fusion of intersecting areas is the emphasis of improvement.

#### REFERENCES

- Marc Soucy, Denis Laurendeau, 1995. A general surface approach to the integration of a set of range views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(4), pp.344-358.
- Zheng Shunyi, Su Guozhong, Zhang Zuxun, 2005, Automatic Reconstruction of 3D Surface Model with 3D Irregular Points Based on Projection Constrain. *Geomatics and Information Science of Wuhan University*, 30(2),pp.154-157.
- Zhang Rongjun, Zhang Zuxun, Zhang Jianqing, 2002, Camera Calibration Using 2D-DLT and Bundle Adjustment with Planar Scenes. *Geomatics and Information Science of Wuhan University*, 27(6), pp.566-571.
- Zhang Zuxun, Zhang Jianqing, 1996. *Digital photogrammetry*. Wuhan Technical University of Surveying and Mapping Press, Wuhan.

Cheng Xiaojun, Zhu Li, Liu Junling, 2005. Three Dimension Modeling Based on Digital Photogrammetry Method. Journal of Tongji University (Natural Science), 33(1), pp.37-41.

Xu Gang, 2006. Image Based Modeling and Rendering. Wuhan University Press, Wuhan.

Feng Wenhao, 2004. Industrial Photogrammetry. Wuhan University Press, Wuhan.

Wang Zhizhuo, 1979. The principle of photogrammetry. Mapping Press, Beijing.

Wolfgang Niem, Marcus Steinmetz, 1996. Camera viewpoint control for the automatic reconstruction of 3D objects. IEEE International Conference on Image Processing, Vol. 3, pp. 655-658.

Jonathan Starck, Adrian Hilton, 2003. Model-based multiple view reconstruction of people. Proceedings of the Ninth IEEE International Conference on Computer Vision, Vol. 2, pp. 915-922.

Zhai Ruifang, Zhang Jianqing, Zheng Shunyi, 2006. Seamless Registration of Multiple Range Image with whole Block Adjustment. First International Multi-Symposiums on Computer and Computational Sciences, Vol. 1, pp. 753-758.

#### **ACKNOWLEDGEMENTS**

I would like to thank the support of Chinese 863 plan (2006AA12Z136).

