GENERATION OF 3D SURFACE MODEL OF COMPLEX OBJECTS BASED ON NON-METRIC CAMERA

Shunyi Zheng, Ruifang Zhai, Zuxun Zhang
syzheng@263.net zhairuifang@hotmail.com zxzhang@supresoft.com.cn
School of remote sensing and information engineering, Wuhan University, Wuhan 430079, China

ABSTRACT

A reconstruction method for complex objects was developed using non-metric images. Firstly, a rotatable platform was designed to capture images and a planar grid board on its top was used for camera calibration. With the help of the platform, images were captured and then processed with a multi-baseline image matching algorithm to produce corresponding point features. The initial external elements of the images can be obtained with traditional photogrammetry approach. These parameters and coordinates of corresponding points were input to block bundle adjustment as initial values, and then point clouds and accurate image parameters were obtained. In the end, a visibility constrained Delaunay triangulation algorithm was used to produce surface models. The methods include artful hardware design, efficient image matching, and rational 3D surface reconstruction. It is low-cost and automatic, and experiment results demonstrate its feasibility and effectiveness.

Index Terms—*Non-metric camera; 3D reconstruction; multi-baseline image matching; vision constrain; Delaunay triangulation*

1. INTRODUCTION

Retrieving three-dimensional geometric information of objects from two-dimensional images is the main research direction in the fields of Photogrammetry, Computer Version and so on; more sophisticated measurement systems have already been explored in fields such as topographic survey. Digital photogrammetry has now been widely applied in surveying and related fields [1]. It can be best exemplified in three-dimensional measurement based on images. Recent years, due to the popularity and generalization of digital cameras, three-dimensional measurement based on images captured by digital cameras now begins to extend to new fields, including ancient building reconstruction and historical relic archiving [5], medicinal reconstruction [8], industrial survey[2], three-dimensional measurement and reconstruction of human faces and bodies[4], engineering survey, etc, and great improvements have been made. However, most measurements are all limited in relatively simple objects, the measurement and three-dimensional expression of complex objects (for instance, human' head, ears and so on) is still a difficult task [1], especially for the simplification and automation of the measurement process. The traditional photogrammetry methods are limited in terms of flexibility [2]. Much research has been done for three-dimensional measurement based on non-metric camera and various methods were proposed in computer vision. But most of these methods concern methodological research and conjugate points were given manually, so these methods are not fully and thus impractical [8]. The results of almost all these methods are the three dimensional point coordinate values of object surface, but the expression of 3D shapes of the objects need surface information.

The paper proposes a feasible and practical algorithm based on non-metric camera for three-dimensional reconstruction of complex objects. Firstly, a rotatable platform was designed to capture images and for camera calibration, and then the images were processed to obtain the geometric information of the object surfaces. Here a multi-baseline image matching algorithm, which is highly reliable and highly accurate, was adopted, and the results were satisfactory. After point clouds were obtained, Delaunay triangulation algorithm based on visibility constrain of the feature points was used to build the surface models of the objects. Experimental results show that this method is feasible for the 3D reconstructions which are the low-cost and automatic.

2. 3D GEOMETRIC INFORMATION ACQUISITION BASED ON IMAGES

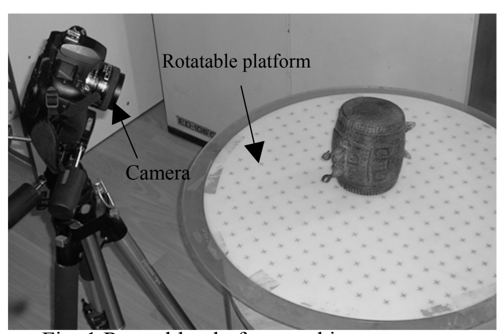

Fig. 1 Rotatable platform and image capture

2.1 Rotatable platform and camera calibration

A rotatable platform (Fig.1) is designed to capture images conveniently. The rotatable platform composes of a planar grid board and a stepper motor which can drive the platform to swivel. The planar grid board is on the top of the rotatable platform, and all three-dimensional coordinates of the grid points marked on the grid board are known. The rotation of the platform and exposal of the digital camera are controlled by a computer, so sequential images can be acquired automatically alone with the rotation of the platform.

Before recovering the 3D information of the object with images, for non-metric camera used in this paper, it is necessary to do calibration to acquire the interior elements (the focus, the principal point of the image and distortion parameters). For this purpose, several images of the planar grid board were captured from different azimuths to do camera calibration with the method proposed in reference[3]. After the calibration, the object was put on the planar grid board, and its images were taken. In order to reduce the difficulties of the image matching, the rotation angle step should be not too long to compose short-baseline image. Next we will discuss it in details.

2.2 Multi-baseline image matching

For images captured by the rotation platform, if the rotation step is too long, the distortion between adjacent images will be large, and the difficulties of the image matching will increase greatly. While if the rotation step is too short, then the baseline between the two adjacent images will be short too, and the intersection angle will be small, this will result in accuracy reduction, especially for the accuracy at deep direction.

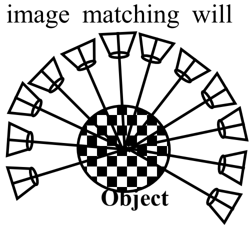

Fig.2 Multi-baseline photography

Taking the two aspects in consideration, the paper proposes a multi-baseline stereo matching strategy which is similar to the principle of multi-view stereo vision in Computer Vision [10]. Firstly, a series of images with short baselines and abundant overlaps were captured. For parallel photography, the distance of the adjacent images should be short, and for convergent photography (Fig. 2), the angle of the adjacent images should be small. Obviously, under this condition, the geometrical distortion among the images will be smaller and image matching becomes much easier, the direct result is reduction of image matching error and increasing of correct ratio of image matching. However, the accuracy of forward intersection will be low for the images acquired in this way. In order to conquer this problem, the paper adopts a multi-baseline matching method which uses multi image points on several adjacent images to do intersection to calculate one object point. In this case, the intersection accuracy can be guaranteed. But it requires the same object point appear on several images. It means that the corresponding points should be transferred between the adjacent images. So the process of the multi-baseline stereo matching is as follows: series of images with short baselines and abundant overlaps were captured, and adjacent images were matched. During the process of image matching, the transfers between the corresponding points should be ensured, i.e. the matched corresponding points in the former pairs were used as the feature points in the next pairs, the corresponding points transfer down until it go beyond the range or the match failed, then the same object point will appear in multi-images to ensure enough intersection angle.

In the traditional technology of single-baseline stereo matching (binocular vision), their forward intersection result is a spatial point without any redundant observation. While using multi-baseline stereo matching method, the corresponding points appear in multi-images, so there would be more redundant observations and it would be easier to detect the match error. As Fig.3 shows, the points in other images are all met at point A expect point a_i in the ist image, the intersection point B of point a_i is apparently the result of wrong match. This algorithm can detect and eliminate the wrong match effectively and can improve the reliability of matching results. Also, it can achieve high accuracy of spatial forward intersection through ensuring enough intersection angles.

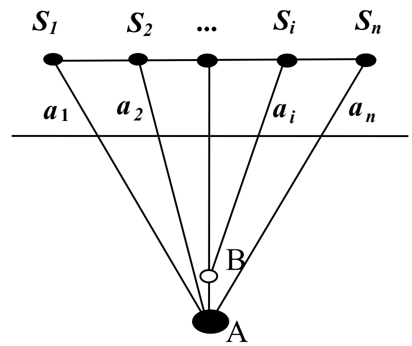

Fig.3 Multi-baseline stereo

The technology of multi-baseline stereo matching can acquire more reliable model point data. The technology suits the surface reconstruction of different objects indoor and outdoor, and the whole process is simple, fast and high-automation.

2.3 Acquisition of accurate 3D geometrical information

After obtaining the corresponding points in the image space with multi-baseline matching approach, what should do next is to determine of 3D point coordinates and the exterior elements of images (positions and attitudes of the camera when capturing the images). Here the traditional aerial photogrammetry can be used as reference. First, the adjacent images are used to construct the stereo pairs, and relative orientations were finished to every pairs with the information of the corresponding points. Then, the models were connected according to the common corresponding points between the stereo pairs, and a free network model was generated, however, as a result of error accumulation, the precision of the model points coordinates is not high. So, on base of that, collinearity equation (show as equation (1)) was used to combine all the images and all the image points to do the adjustment, and it is the free network bundle adjustment in photogrammetry.

$$x - x_0 = -f\frac{a_1(X - X_S) + b_1(Y - Y_S) + c_1(Z - Z_S)}{a_3(X - X_S) + b_3(Y - Y_S) + c_3(Z - Z_S)} \quad (1)$$
$$y - y_0 = -f\frac{b_1(X - X_S) + b_2(Y - Y_S) + c_2(Z - Z_S)}{a_3(X - X_S) + b_3(Y - Y_S) + c_3(Z - Z_S)}$$

The equation expresses the geometrical relationship among the image points, the object points and the center of the camera exactly, and they are all on the photography ray. In the equation, (x, y) are the coordinates of image point, x_0, y_0 are the position of the camera principal point of photograph, f is the focus of the camera, X_S, Y_S, Z_S are the position parameters of exterior elements, a_i, b_i, c_i are the nine direction cosines which is also called as rotation matrix deriving from the three attitude angles. Error equation according to the collinearity equation is as following [9],

$$v_x = \frac{\partial x}{\partial X_s}\Delta X_s + \frac{\partial x}{\partial Y_s}\Delta Y_s + \frac{\partial x}{\partial Z_s}\Delta Z_s + \frac{\partial x}{\partial \phi}\Delta\phi + \frac{\partial x}{\partial \omega}\Delta\omega + \frac{\partial x}{\partial \kappa}\Delta\kappa + \frac{\partial x}{\partial X}\Delta X + \frac{\partial x}{\partial Y}\Delta Y + \frac{\partial x}{\partial Z}\Delta Z - l_x \quad (2)$$
$$v_y = \frac{\partial y}{\partial X_s}\Delta X_s + \frac{\partial y}{\partial Y_s}\Delta Y_s + \frac{\partial y}{\partial Z_s}\Delta Z_s + \frac{\partial y}{\partial \phi}\Delta\phi + \frac{\partial y}{\partial \omega}\Delta\omega + \frac{\partial y}{\partial \kappa}\Delta\kappa + \frac{\partial y}{\partial X}\Delta X + \frac{\partial y}{\partial Y}\Delta Y + \frac{\partial y}{\partial Z}\Delta Z - l_y$$

Where v_x and v_y are the correction of the observation, $\Delta X_S, \Delta Y_S, \Delta Z_S, \Delta\varphi, \Delta\omega, \Delta\kappa$ are the correction of unknown parameters, $\Delta X, \Delta Y, \Delta Z$ are the corrections of object point coordinates. The initial values of the adjustment model X、Y、Z are the results of forward intersection with the initial values of the elements of orientation gained after the relative orientation and connection of models.

If there are control points, their information can be added, and the model is built under the coordinate system of control points. In the case of no control points, the coordinate system of the first image can be appointed as the global coordinate system and then the model can be built under such global coordinate system. As a result of not importing the control points, it is a rank defect free network. So it needs to introduce coordinate benchmarks to solve the problem. The length of benchmarks can be defined as the distance between the first image and the second image or the longer X-component of the baseline between the first image and the second image, usually the latter definition is used. After bundle adjustment, the 3D spatial coordinates and the elements of exterior orientation of each image were obtained.

3. SURFACE MODEL GENERATION

With the method introduced above, the 3D point clouds of the object surface can be acquired. To present the 3D shape of the objects, not only a set of feature points are needed, but also the inter relations among these points. Building the surface model automatically according to the 3D spatial points is a hot point in Computer Graphic and Computer Version and so on. Although many scholars have proposed lots of theories and algorithms, there still many problems are waiting to be settled. In this section a visibility constrain based Delaunay triangulation algorithm was proposed. The 3D surface model of the object can be constructed with this method according the information whether the feature points are visible in the image.

3.1 3D Delaunay triangulation

It is well known, the Delaunay triangulation of a point set S consists of tetrahedrons formed by points in S whose circumscribing spheres enclose on other points in S. The set of tetrahedron covers the convex hull of S[11].

After Delaunay triangulation, we can get the smallest convex hull enveloping the point set. However common objects to be reconstructed are not convex, so until now, the above methods can not exactly reconstruct the surface model of objects.

3.2 Visibility constrain

For the 3D spatial points computed from several images, the projection of these points can be found at the corresponding images in which they are visible. If the connecting lines between the view point and these points intersect with the some existed tetrahedrons, then these tetrahedrons should be deleted. The reason is that if these tetrahedrons exist, then the points behind them will be invisible (see Fig.4). So for every image, the view point and the visible points in the image were connected with lines, and then the tetrahedrons intersect with these lines was detected and deleted. As Fig.4 shows, if point A is visible from the view point V, the tetrahedron ABCD intersecting with line *VA* should not exist, and it should be deleted. After deleting the tetrahedron which should be deleted and recording the most outside face of the remained tetrahedrons, we can get a surface model which is the exact surface model of the object.

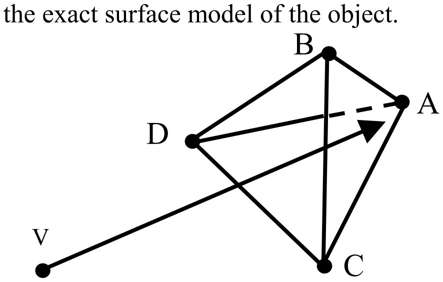

Fig.4 Visibility constrain used in DELAUNAY triangulation

4. EXPERIMENT AND ANALYSIS

A hardware device and corresponding software system were designed based on the principle and the algorithms described above. In order to prove the effectiveness of the approach, experiment was made on a bronze vessel with a complex surface on which there is something like two gluttonous birds. This bronze vessel is a Chinese relic of about 3000 years ago. The camera used in the experiment is Kodak DCS Pro SLR/n, whose geometric resolution is 4500*3000 pixels and focus lens is 50mm.

Firstly, camera was calibrated with rotatable platform. The control field used to do calibration is the planar grid on the rotatable platform which has 326 grid points and the grid interval is 30mm. Its spatial coordinates have been measured beforehand. Then images of bronze vessel were captured every 10 degrees while the rotatable platform was controlled to rotate. Totally 36 images were captured, shown in Fig.5. Secondly, the corresponding points on the surface of the objects were matched through multi-baseline image matching and then used to calculate the initial values of the exterior parameters of the images and 3D coordinate of corresponding points are computed through intersection. All these information was input into bundle adjustment program to compute the accurate solution of 3D coordinates of corresponding points and exterior parameter of each image. Then visibility constrain based Delaunay triangulation algorithm was used to produce the surface model. Initial result of Delaunay triangulation is showed as (Fig. 7a), the result with visibility constrain is showed as (Fig. 7b), the surface model with texture showed as (Fig. 7c).

After eliminating all the points with 2 overlaps (the feature points appear only in two images), totally 112533 feature points were remained (Fig.6), of which 83487 points appear in three-four images and 29064 points appear in more than five images. The smallest intersection angle is 6.458 degrees, the biggest one is 119.675 degrees, and the average angle is 29.855 degrees.

The results show that visibility constrain based Delaunay triangulation can produce the true shape of the objects exactly. But it can't ensure that all the inexistent tetrahedrons are deleted. General speaking, more selected points and more images, the higher possibility that these inexistent tetrahedrons are deleted. At last few inexistent triangular should be deleted manually.

5. CONCLUSIONS

This paper presents research on three-dimensional reconstruction of complex objects based on non-metric camera. A whole set of solution including hardware and software was proposed, and experiment results demonstrate its effectiveness. The algorithms can retrieve three-dimensional surface models of complex objects. It is low-cost, automatic and can be used in broad fields, such as digital protection of ancient historical relics, detecting industrial accessory, and so on. further work is needed to improve the algorithm and explore its commercialization and practicability by testing more objects with different shapes.

Fig. 5 Original image

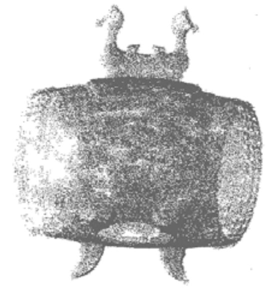

Fig. 6 Retrieved point cloud data

a) Initial surface model b) Surface model after using visibility constrain c) Surface model with texture

Fig. 7 Results of 3D reconstruction

6. REFERENCES

[1] S. Zheng, G. Su, and Z. Zhang, Automatic Reconstruction of 3D Surface Model with 3D Irregular Points Based on Projection Constrain. *Geomatics and Information Science of Wuhan University*, 2005, 30(2)：154-157.

[2] Feng, W., *Industrial Photogrammetry*. WUHAN: Wuhan University Press, 2004.

[3] Zhang, Y., Zhang, Z., Zhang J.: Camera Calibration Using 2D-DLT and Bundle Adjustment with Planar Scenes. Geomatics and Information Science of Wuhan University, 2002, 27(6):566-571.

[4] Nicola, D.A.: Human face modeling from images. Oral Speech in 3rd International Image Sensing Seminar on New Development in Digital Photogrammetry. Japan: Gifu, 2001:28-29

[5] Pollefeys, M.: Self-calibration and Metric 3D Reconstruction from Uncalibrated Image Sequences. Belgium: Katholieke University, 1999

[6] Zhang Z., Zhang, J.: Digital Photogrammetry. WUHAN: Wuhan Technical University of Surveying and Mapping Press, 1996.

[7] Wang Z.: Principle of Photogrammetry. BEIJING:Surverying and Mapping Press, 1979.

[8] Xu, G.: Image based modeling and rendering [M]. Tokyo: KINDAI KAGAKU SHA CO.,LTD, 2001.

[9] Deng, F.: Research on LIDAR and Digital Images Registration and Objects Extraction. Wuhan: Wuhan University, 2006.

[10] Zhang, Y.: Image Understanding and Computer Vision. BEIJING: Tsinghua University Press, 2000.

[11] Joe, B.: Three-Dimensional Triangulations from Local Transformations. SIAM J. Sci. Statist. Comput. 10: 718-741, 1989

7. ACKNOWLEDGEMENT

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