3-D Reconstruction of Real Objects Using an Android Device

Krzysztof Kowalak[1], Łukasz Kamiński[1], Paweł Gardziński[1], Sławomir Maćkowiak[1], and Radosław Hofman[2]

[1] Poznań University of Technology, Poland
{kkowalak,lkaminski,pgardzinski,smack}@multimedia.edu.pl
[2] 3D Vision Sp. z o.o., Poland radekh@3d-vision.com.pl

Summary. In the paper, the autonomous system of reconstruction of 3-D model based on the matching characteristic features between the images for mobile devices with Android OS is proposed. Our method focuses on fully automated system with marker less calibration method. The experimental results show that although the reconstructed objects contain certain artifacts or loss, the end result can be successfully used by the average user.

1 Introduction

In recent years 3-D printing, 3-D visualization of a scene in everyday life and a video content creating have earned massive popularity. A large number of 3-D models is generated, put together in one scene, allowing viewers to navigate through it or print it and used it in prototyping process. People in the industry spent uncountable hours trying to model the world with detailed environment. This can be troublesome for most people. It would be nice if anyone can reconstruct any scene depicted in their photographs in a fairly accurate 3-D model.

This paper presents innovative approach towards 3-D reconstruction of objects, which is characterized by low-cost of the system. The solution is based on mobile devices with Android operating system as an image acquisition system and an application which performs image processing operations and is running eg. in the cloud systems.

The most popular methods of the 3-D model objects reconstruction based on images are: techniques based on the estimation of depth maps from two or more views of the object [7,8], reconstructions based on the matching characteristic features between the images [9,10] and volumetric reconstruction techniques from multiple camera views [11,12].

The techniques based on the estimation of depth maps from many images are very complicated and time consuming. This process can be accelerated by the application of ToF cameras or cameras using structured light (eg. an infrared light KinectOne) and generate directly a cloud of points which is the distance to the object (the depth of the scene). During the conducted research on the reconstruction of 3-D models of objects using volumetric reconstruction algorithm, the authors came across a problem with the accuracy of the internal and external camera parameters and the effect of these parameters on the quality of the estimated 3-D model. These camera parameters have a significant impact on the quality of the estimated model. Proposed solution uses characteristic feature points matching between two images to perform 3-D reconstruction. Authors decided to use marker less calibration method. This approach is dedicated for most people who want to create the 3-D models via mobile devices by creating the series of the photos of the object and then send these images to the server and nothing more. The details are presented in the next section.

This paper is organized as follows: next section describes the proposed system of the 3-D model reconstruction and the advances operations on images. Section 3 contains results obtained by running these apps. The Section 4 contains the conclusions and the future scope of this work.

2 System description

Proposed system of 3-D model reconstruction is using an Android device. Scanned object must be captured from different sides in order to build a 3-D model. Captured images are used to reconstruct the spatial information about the scene. General diagram of proposed system is shown on Figure 1.

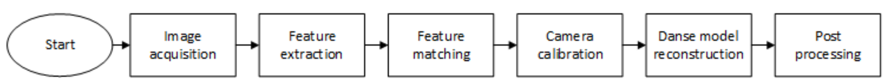

Fig. 1. General diagram of the proposed system.

2.1 Image acquisition

Image acquisition of scanned object is done using a specially prepared Android application. The main purpose of application is to support the user during scanning process. In order to build a model, user should take a series of images of object from different angles. It is worth mentioning that there is no calibration process of Android device. Intrinsic and extrinsic parameters of camera and lens distortion coefficients are estimated in the later stages of the algorithm. Besides the images of object, designed application also provides information about model and manufacturer of the device as well as focal length.

2.2 Feature extraction

Spatial information about observing scene must be known in order to determinate camera coordinates for each image in one common coordinate system. Many camera calibration methods are well known in literature [1]. Generally, these methods can by divided into two groups. One group of them is using marker with known geometry, and the second group is marker less. Methods which are using a marker are less complex than marker less methods, and are giving accurate results. Unfortunately, the need to use marker carries some complications. The end user would have to have a marker in possession and know how to use it. Therefore authors decided to use marker less calibration method. In order to calibrate the camera without a marker it is needed to find some points of interest in all views. In proposed system SIFT[2] feature points are used. SIFT points are resistant and remain stable in case of scale, orientation and lighting change. Additionally for each SIFT point his own descriptor is calculated which is used in next step of algorithm.

2.3 Feature matching

The next step of the algorithm is feature points matching between two views. Individual views are matched in each to each scheme. This is more complex solution than chain matching but provides resistance to interruption in case of impossibility of matching of adjacent pair of images. Feature point matching relies on descriptors comparison of these points. Brute Force method with L2 distance metric is used to point matching between two views. The result of this operation is a list of matched pairs among two images. It may happen that some points are matched incorrectly. For this reason the correction step is performed which is using a fundamental matrix. Fundamental Matrix is calculated between two images in RANSAC [3] scheme and describes the transformation of point set from first to second image.

2.4 Camera calibration

The main goal of this step is to determine camera orientation and position in space for each image of scanned object. Calculated intrinsic and extrinsic parameters must be accurate that means reprojection error of 3-D point into image plane need to be as small as possible. The problem of camera parameters determination can be formulated as non-linear minimization of mean squared error and can be solved by Levenberg - Marquadt [4] algorithm.

The camera parameters estimation process starts with single pair of images. The initial pair should have high number of matched feature points and proper displacement so that the 3-D position of points can be identified well. Initialization of new view starts with using a DLT[5] (Direct Linear Transform) algorithm and RANSAC[3]. After determining the parameters of the camera, points which are observed by this camera are added to 3-D model of object. Example of sparse model reconstructed in this step is shown in Figure 2.

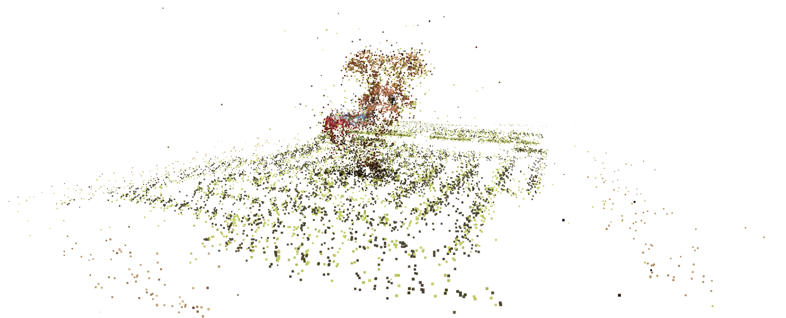

Fig. 2. Example of reconstructed sparse model.

2.5 Dense model reconstruction

Dense model algorithm relies on usage of SIFT characteristic features and projection matrices designated from previous step and increase model coverage by patches. Each characteristic point is a central point in patch with μ x μ size, where μ is equal 5 or 7. For single patch normal vector $n(p)$ is assigned, except central point $c(p)$. Normal vector is always directed on outside of the model. The next step is patch expansion. New patches are created in empty space between neighboring patches. When the expansion is finished, patches filtering process is performed. For each view, location of each patch is verifying. If some of them is in front of or behind set of patches plane, it is removed. Based on this dense model, surface reconstruction can be made.

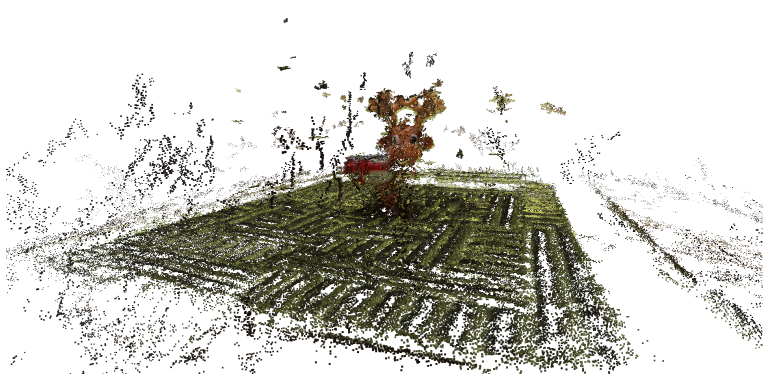

Fig. 3. Dense model reconstruction.

2.6 Post processing

The dense model which was obtained in previous step need some post processing. First of all in reconstructed model appear some points which do not belong to the scanned object. Additionally to build a complete 3-D model there is need to perform surface reconstruction step. Point cloud filtering is done by SOR [6] (Statistical Outlier Removal) filter. This filter is based on stochastic analysis of all points in the cloud and rejection of those which do not meet certain criteria. During the filtration process the distance between point and its neighborhood is calculated. Assuming that the distribution of these distance is Gaussian, point that do not match the distribution are removed. Example of point cloud before and after the filtration process is shown in Figure 4.

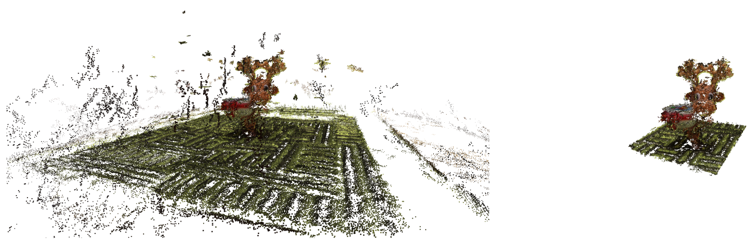

Fig. 4. Example of SOR filter application.

Last step of algorithm is surface reconstruction. Surface reconstruction is complex problem because the quality of reconstructed point cloud is mostly low due to noise. In order to perform surface reconstruction Poisson [7] surface reconstruction algorithm was used. This algorithm treats model reconstruction problem as spatial Poisson problem. This method is resistant to noise in point cloud because is considering all points simultaneously while giving satisfactory result. The result of this method is shown of Figure 5.

3 Experimental results

Studies conducted in this paper concerned the estimate subjective quality of reconstruction on the basis of prepared test sequences of real objects.

Data for the tests were recorded using a popular smartphone with Android OS. Four image sequences which vary in terms of difficulty of reconstruction were recorded. Two test sequences have low complexity and two others have a higher degree of complexity. All sequences were prepared in two versions: as a series of images and a video sequence.

Moreover, all sequences were prepared in different lengths, in order to verify the effectiveness of the reconstruction where there are plenty of images

Fig. 5. Example of reconstructed surface from point cloud.

around the object (high density scan), and in the case of the small number of input images.

According to project assumptions, the solution presented in this paper is prepared for the end user who does not have the necessary knowledge to handle most of the currently available solutions. As already described in section 2, the algorithm does not require any configuration by the user. In connection with this assumption, widely available data sets do not meet the prescribed requirements. These sequences are not recorded from the hand as it probably will do the average user. It was therefore necessary to record our own sequences that satisfy the above requirements.

The effectiveness of the reconstruction in the presented solution was evaluated based on a subjective assessment of quality of reconstruction by comparing the input images with model reprojection to the image plane. The experimental results of proposed approach is shown on Figure 6.

The developed system of autonomous reconstruction of the 3-D model gives very good results. Although the reconstructed objects contain certain artifacts or loss, the end result can be successfully used by the average user. The reconstructed objects can easily be used to further work in graphic programs or 3-D printers. Computation time depends on number of images and images resolution. The presented technique is characterized by wide application, but needed further study in order to further improve the quality of reconstruction and reducing computation time.

4 Conclusions

In this paper, the novel system for 3-D model reconstruction is proposed. The proposed system is using only an Android device in order to capture images of scanned object. Special application for Android was designed to guide the user during scanning process. All calculations take place on the server where images were uploaded by application. The main advantage of

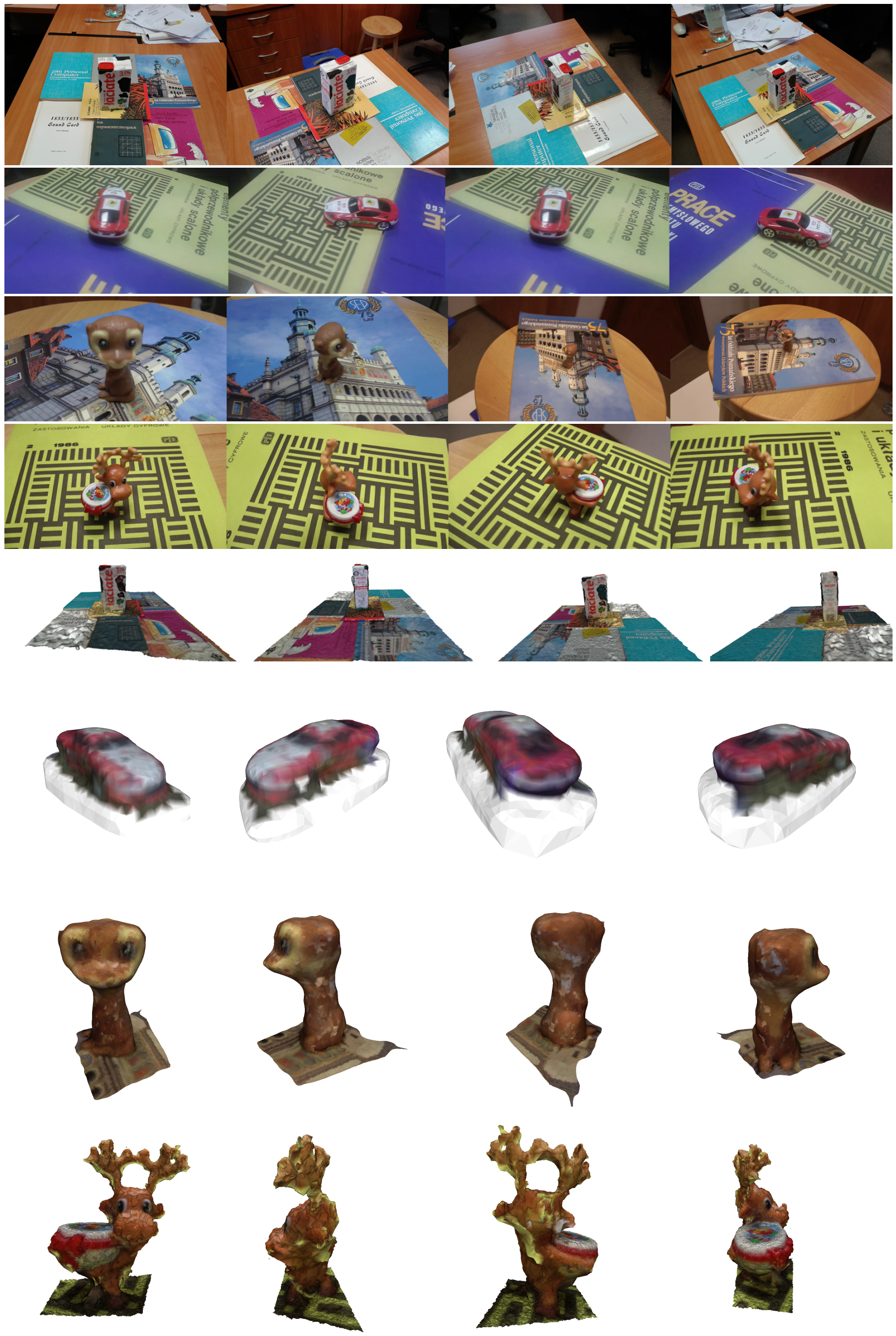

Fig. 6. Experimental results of the 3-D reconstruction of the test objects.

proposed approach is that user does not need to have any special knowledge about scanning procedure. The marker is also unnecessary.

At the end of the algorithm the user receives complete 3-D model of scanned object. The experimental results shows that the quality of 3-D model is satisfactory. The model is saved in PLY file format which can be used by most graphics programs available on the market. Reconstructed model of object can be further processed or printed on 3-D printer.

Acknowledgement

The project *Innovative stereoscopic webcam with dedicated software for image processing prototype construction* is the project co-founded by European Union (project number: UDA-POIG.01.04.00-30-018/10-03, realization period: 02.2012-06.2014). European funds for the development of innovative economy.

References

1. Cyganek B (2013) Object Detection and Recognition in Digital Images: Theory and Practice, Wiley
2. Lowe David G. (1997) Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, 60, 2:91-110
3. Fischler M. , Bolles R. (1987) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Readings in computer vision: issues, problems, principles and paradigrims, 726 740
4. Nocedal J. , Wright S. (1999) Numerical Optimization, Springer Verlag, New York
5. Hartley R. , Zisserman A. (2003) Multiple view geometry in computer vision, Cambridge University Press
6. Point Cloud Library (PCL) Documentation http://dev.pointclouds.org/projects/pcl/wiki
7. Seitz S.M. , Curless B. , Diebel J. , Scharstein D. , Szeliski R. (2006) A Comparison and Evaluation of Multi-View Stereo Re-construction Algorithms. In: IEEE Conf. on Computer Vision and Pattern Recognition, pp. 519-528
8. Xiao J. , Chen J. , Yeung D.-Y. and Quan L. , (2008) Learning twoview stereo matching, In Proceedings of the 10th Eu-ropean Conference on Computer Vision: Part III, ECCV 08, Springer-Verlag, Berlin, Heidelberg, pp. 15-27
9. Snavely N. , Seitz S. and Szeliski R. , (2006) Photo tourism: ex-ploring photo collections in 3d, In ACM Transactions on Graphics (TOG), volume 25, ACM, pp. 835-846
10. Snavely N. , Seitz S. and Szeliski R. , (2008) Modeling the world from internet photo collections", International Journal of Computer Vision, 80(2): pp. 189-210
11. Goldlcke B., Magnor M., (2003) Real-time, free-viewpoint video rendering from volumetric geometry, In Proc. Visual Compu-tation and Image Processing (VCIP), Lugano, Switzerland, 5150(2): pp. 1152-1158
12. Baker H., (1977) Three-dimensional modelling, In Proc. of the 5th International Joint Conference on Artificial Intelligence, Cambridge, MA, USA, pp. 649-655