MŰEGYETEM 1782
Budapest University of Technology and Economics
Faculty of Civil Engineering
Department of Photogrammetry and Geoinformatics

Modeling based on depth images and non-metric imagery

Theses of the PhD Dissertation

Bence Attila MOLNÁR

M.Sc. in Surveying and Geoinformatics

Supervisor:
Dr. Ákos DETREKŐI, Dr. Árpád BARSI

Budapest, 2013

1. Objective

Our everyday devices are equipped with sensors also suitable for spatial data acquisitions. High distribution of sensors is indicated by two synergistic effects: the needs for owning a popular device and the low price.

If inexpensive sensors are available, it raises the possibility of using them for precise modeling purposes. In my thesis I investigated the ability and accuracy of commercial and mass produced devices for object recognition and reconstruction. A cost-effective way will help meet the growing demand of modeling, which occur in several areas and spatial measurements in different application fields, such as robotics or medical sciences. As the capacity of computation and data storage developed, it become more common to collect mass data rather than extremely accurate data. Accuracy of the applications are increased in a statistical way to the demanded level based on the redundant measurements.

2. Web based photogrammetric application

A new era is coming in photogrammetry by the commercial digital cameras and smart phones equipped with cameras. The low-end devices are available for a low price therefore they are widely used. If there is affordable photogrammetric application and users had some photogrammetric knowledge, low-end cameras will perfectly support 3D modeling. Obviously the model accuracy will be lower, however, it can be increased e.g. by using repeated, multiple images.

Nowadays web applications are used in our everyday life. The major advantage is that we do not need to install any software on our own machine, and we can continue our work on any computer on the network. Users don’t need to worry about storage issues; data are usually stored in cloud servers. These opportunities yield to a widespread use.

An optimized database and transmission system should be designed during the development phase of a web based application. Photo processing applications deal with high amount of data, therefore bandwidth optimized communication is particularly important; images are transmitted in tiles and stored at the client (cache) and on the server as well. User experiences slow performance if the communication is poorly designed, however, data should be saved real-time.

A web based photogrammetry application was developed to make photogrammetry available for everyone in an easy way. Server-client communication is optimized for bandwidth, the efficient calculation and data storage is performed on the server and users can use a thin client to perform 3D reconstruction.

3. Robust photogrammetric object reconstruction

High amount of users require robust operation of the processing programs. In addition to the didactic user-friendly interface, calculation procedure has to be robust, therefore an algorithm was developed which is capable of processing 3D reconstruction for poorly designed photogrammetric networks. To make the procedure suitable for all camera types, Direct Linear Transformation (DLT) is used for processing and Least Squares Method is used for optimization (Detrekői, 1991). The permutation of poorly conditioned matrixes is high in photogrammetry, therefore Singular Values Decomposition (SVD) is applied (Eckart and Young, 1938). Considering that the Least Squares Method is very sensitive on outliers that are made most likely by inexperienced users, Huber method (Huber, 1981) was applied to increase robustness.

4. MS Kinect depth camera

Depth images are recorded by depth cameras; all pixel represents a distance, reflected from the object in the field of view. Depth cameras use light for measurement and can achieve up to 30 fps sampling rate. Formerly, depth cameras had low accuracy, high noise level and high costs therefore they were not widely used. MS Kinect (Figure 1) brought a new era as it is a low cost depth camera for gaming consoles.

Figure 1. MS Kinect

Kinect captures 640x480 px resolution images by both RGB and depth camera, with 30fps. By transforming the depth images to a point cloud and co-registering the RGB and depth camera (Figure 2), a colored point cloud is available in real-time.

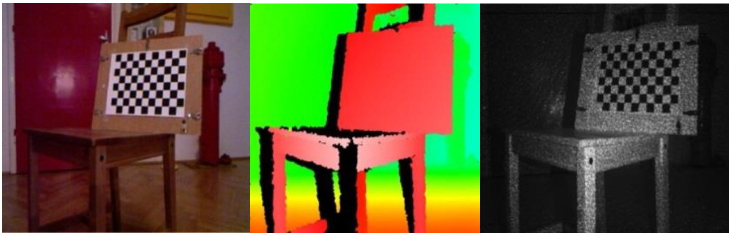


Figure 2. Kinect raw data: RGB, depth and intensity images

Depth values are rounded by a non-linear function on the depth image (Macknojia *et al.*, 2012) that results in a leveled point cloud with layers parallel to the image plane (Figure 3), the level spacing increases by the object distance; over a 3 meter range the spacing is more than 4 cm.

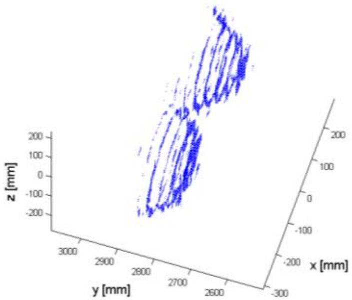


Figure 3. Quantized point cloud

Common fitting methods for curves and surfaces are performed on point basis by minimizing object point distance; which method doesn't perform well if a non-planar object is fitted on a quantized point cloud. Fitting on a sphere yields to a downscaling since higher amount of points can be found inside due to the rounding. In other words, the mean circle of a level doesn't fit to the surface (Figure 4); this is a systematic error and it highly depends on the current constellation of levels and sphere locations.

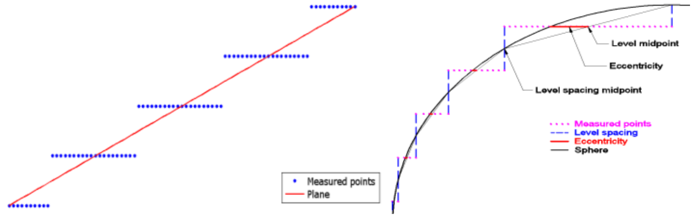


Figure 4. Cross section of a plane and sphere gathered by a depth camera

5. Sphere fitting on quantized point

Midpoints of the quantization level steps form concentric circles; these circles are on the surface. Spheres can be fitted without systematic errors on these circles; therefore sphere fitting can be solved on quantized point cloud (Figure 5).

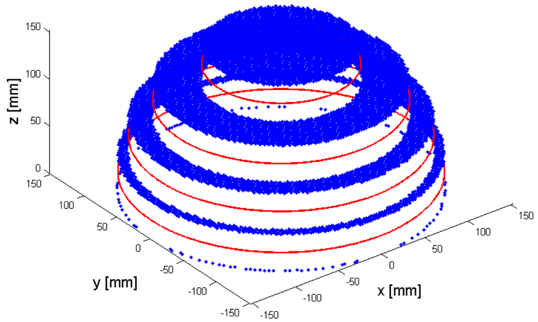


Figure 5. Quantized point cloud captured from a sphere and level midpoint circles

As depth cameras' primary results are depth images, standard image processing algorithm was used to detect level steps; circles were detected by Canny edge detection (Canny, 1986) and Hough transformation (Hough, 1959). Depth values of the detected circles were calculated as the mean values of neighboring levels' depth values; afterwards the modified depth image was transformed to the spatial coordinate system. The developed sphere fitting method (called KvantFit) was compared to two different point error based sphere fitting methods (Figure 6).

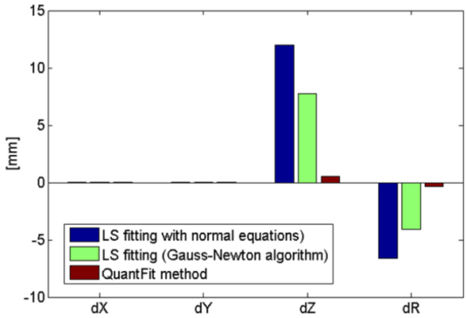


Figure 6. Comparison of common and KvantFit sphere fitting methods

Results based on simulated data prove the applicability of the developed algorithm; downscaling effect was eliminated and depth positioning accuracy was increased. The developed algorithm was also tested on real data set; data was repeatedly gathered by increasing object distance and sphere fitting was applied with both common and KvantFit algorithm. The outcome functions shows the same results as the simulated data and function noise is also reduced (Figure 7).

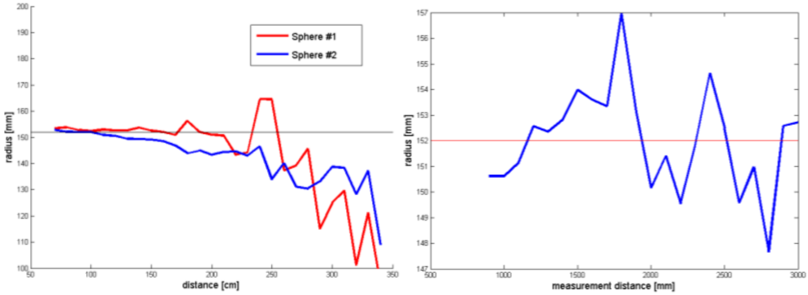


Figure 7.Common and KvantFit sphere fitting in function of object distance

6. Accuracy test for depth cameras

No commonly used comprehensive accuracy test for depth cameras exists. Despite the fact that there were accuracy tests made before; those only focused on a single property such as point cloud comparison or point based STD (Chow, 2011; Khoselham, 2011). A new test sequence was created for investigating repeatability, relative and absolute accuracy; the test was applied on MS Kinect. In general, accuracy of depth cameras highly depends on object distance, therefore the tests was performed on the entire measurement range with 10 cm steps; yielding to a complete accuracy function.

The first test was carried out under the same conditions allowing to calculate the STD for each pixel. The second test is about plane fitting on point cloud gathered from planar surface, afterwards the fitting residuals were compared. To give more information on relative spatial accuracy, sphere detection was applied on point cloud and was compared to Faro terrestrial laser scanner (TLS) data. Finally, sensor and sphere positions were validated by TLS data to check absolute accuracy.

Test	STD [mm]	STD (percentage of the distance) [%]
Repeatability	5	0.25
Plane fitting	7	0.3
Relative accuracy test	3.5	0.2
Absolute accuracy test	0.75	0.1

7. Automatic sphere detection

The depth cameras are also capture the background and surrounding area of the object. Therefore the point cloud filtering and object recognition is very important; control points should also be detected. An automatic algorithm was developed to process depth images and detecting spheres; the algorithm is robust, works perfectly for any size of sphere and at any object distance. Canny edge detection and Hough transformation algorithms were applied to process depth images; the detected spheres on depth images became ready for a high accuracy sphere fitting by KvantFit algorithm.

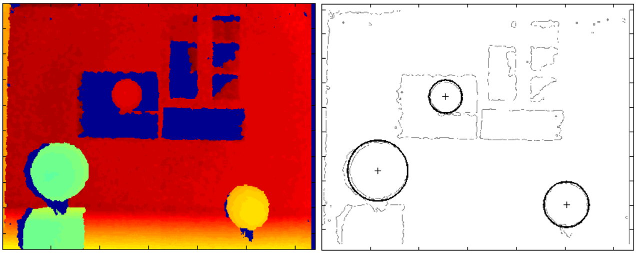
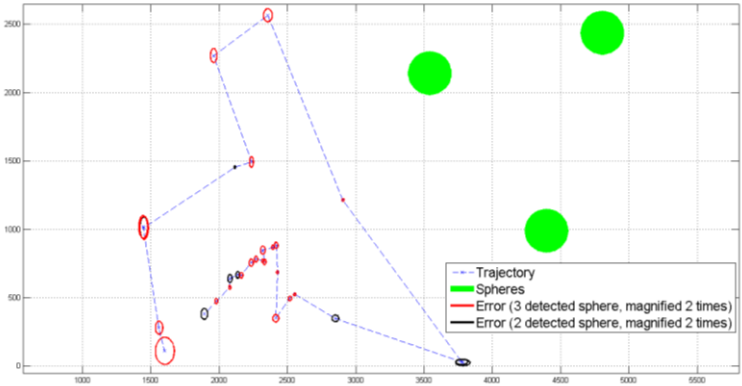


Figure 8. Raw depth image with spheres and automatic detection result

8. Trajectory reconstruction

Once the control point's accurate coordinates in the reference system are known the transformation parameters of the sensor and reference coordinate system can be determined. On one hand it allows the accurate aligning of point cloud sequences; on the other hand the trajectory can be reconstructed if the sensor moves.

An algorithm was developed for sensor position calculation with least squares method and resection. Main goal is an indoor navigation therefore some transformation constraint was applied: 3 different direction shifts and just rotation around the vertical axis are possible.



9. ábra Helyreállított mozgási trajektória

9. Theses

Thesis 1.

To effectively handle ill-conditioned photogrammetric equation systems resulted by imperfect image capturing constellations, I extended the Direct Linear Transformation (DLT) method for object reconstruction by robust Huber estimation and Singular Value Decomposition (SVD).

Related publications: Molnár, 2010b; Molnár, 2010c; Molnár, 2010d; Molnár, 2010e

Thesis 2.

I developed cloud based photogrammetry application. Users are able to perform photogrammetric data processing using efficient computing resources and storage by a thin client. Design and development of server client communication was optimized for bandwidth.

Related publications: Molnár, 2010a; Molnár, 2010b; Molnár, 2010d; Molnár, 2010e

Thesis 3.

I developed a quantization parameter based sphere fitting method to reduce the radial regular error on point clouds provided by depth camera; the applicability of the developed method was tested on simulated and real data.

Related publications: Molnár *et al.*, 2012a; Molnár és Toth, 2013a; Molnár és Toth, 2013b

Thesis 4.

I developed an accuracy analysis procedure to qualify depth cameras that enable to derive relative and absolute accuracy values. Applying tests on Microsoft Kinect depth sensor proved that Kinect can be used with 1 cm standard deviation at less than 3 meters object distance.

Related publications: Molnár *et al.*, 2012a; Molnár és Toth, 2013a; Molnár és Toth, 2013b; Toth *et al.*, 2012

Thesis 5.

I developed continuous indoor navigation solution for depth cameras based on automatic depth image filtering and processing. The developed method is applicable for real-time trajectory reconstruction with centimeter accuracy.

Related publications: Molnár és Toth, 2013a; Molnár és Toth, 2013b; Molnár *et al.*, 2012b

10. Publications on the subject of Theses

1. Molnár, B. (2010a): Developing a web based photogrammetry software using DLT, Pollack Periodica Vol. V Number 2/Aug 2010, pp. 49-56.
2. Molnár, B. (2010b): Direct Linear Transformation Based Photogrammetry Software on the WEB, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII Part 5, pp. 462-465.
3. Molnár, B. (2010c): Web alapú fotogrammetriai alkalmazás pontossági vizsgálata, Geomatikai Közlemények, XIII/2 kötet, pp. 101-106.
4. Molnár, B. (2010d): Robosztus becslést és DLT-t alkalmazó web alapú fotogrammetriai alkalmazás fejlesztése, Geomatikai Közlemények, XIII/1 kötet, pp. 91-95.
5. Molnár, B. (2010e): Robosztus becslések használata web alapú fotogrammetriai kiértékeléshez, Geodézia és Kartográfia, LXII kötet 1. szám, pp. 24-27.
6. Molnár, B.; Toth, C. K.; Detrekői, Á. (2012a): Accuracy test of Microsoft Kinect for human morphologic measurements, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIX-B3, pp. 543-547.
7. Molnár, B.; Toth, C. K.; Grejner-Brzezinska, A. D. (2012b): Sphere fitting on MS Kinect point cloud, MAPPS/ASPRS 2012 Specialty Conference, Tampa, Amerikai Egyesült Államok. In print.
8. Molnár, B.; Toth, C. K. (2013a): MS Kinect - Flash LiDAR pontossági vizsgálat és gömbillesztés, Geomatikai Közlemények, XVI/1 kötet, In print.

9. Molnár, B.; Toth, C. K. (2013b): Spherical Target Based Trajectory Recovery From Kinect Depth Imagery, ASPRS 2013 Annual Conference, Baltimore, Amerikai Egyesült Államok, 2013.03.24-2013.03.28.
10. Toth, K. C.; Molnár, B.; Zaydak, A.; Grejner-Brzezinska, A. D (2012): Calibrating the MS Kinect Sensor, ASPRS 2012 Annual Conference, Sacramento, Amerikai Egyesült Államok, pp. 538-546.

11. References

1. Canny, J. (1986): A Computational Approach to Edge Detection, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. VIII Part 6, pp. 679-698.
2. Chow, J. C. K.; Ang, K. D.; Lichti, D. D.; Teskey, W. F (2012): Performance Analysis of Low-Cost Triangulation-Based 3D Camera: Microsoft Kinect System, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIX-B5, pp. 175-180.
3. Detrekői, Á. (1991): Kiegyenlítő számítások, Tankönyvkiadó, Budapest, p. 685.
4. Eckart, C., Young G. (1936): The approximation of one matrix by another of lower rank, Psychometrika, Vol. I Part 3, pp. 211-218.
5. Hough, P.V.C. (1959): Machine Analysis of Bubble Chamber Pictures, Proc. Int. Conf. High Energy Accelerators and Instrumentation, Vol. VII, pp. 554-556.
6. Huber, P. J. (1981): Robust Statistics, John Wiley & Sons, New York, p. 312.
7. Khoshelham, K. (2011): Accuracy Analysis of Kinect Depth Data, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII -5/W12, pp. 133-138.
8. Macknojia, R.; Chávez-Aragó, A.; Payeur, P.; Laganière, R. (2012): Experimental Characterization of Two Generations of Kinect's Depth Sensors, IEEE International Symposium on Robotic and Sensors Environments, pp. 150-155.