БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ • BULGARIAN ACADEMY OF SCIENCES

ПРОБЛЕМИ НА ТЕХНИЧЕСКАТА КИБЕРНЕТИКА И РОБОТИКАТА, **70**
PROBLEMS OF ENGINEERING CYBERNETICS AND ROBOTICS, **70**

София • 2018 • Sofia

3D modelling for object recognition with depth sensors

D. Chikurtev[1], I. Rangelov[2], N. Chivarov[2], D. Karastoyanov[1]

[1]Institute of Information and Communication Technologies – BAS

[2]European Polytechnical University - Pernik

Email: dchikurtev@gmail.com,
rangelov.ivaylo@yahoo.com, **nshivarov@code.bg**, **dkarast@iinf.bas.bg**

Abstract: *In this paper, we represent methods for 3D modelling of objects. These methods are used for object recognition in the field of service robotics. Presented methods are manual scan method by hand and manual scan method by a robot. In the paper are described 3D sensors - Kinect and Intel RealSense and 3D software for scanning and processing. For conducting experiments of object scanning are applied the proposed methods using the mobile robot and Microsoft Kinect for Windows. In addition, the result shows the advantages and disadvantages of the methods.*

Keywords: *3D sensor, 3D scan, object recognition, Kinect, Intel RealSense, mobile robot*

1. Introduction

Computer vision is a rapidly growing field devoted to analyzing, modifying, and high-level understanding of images. Its objective is to determine what is happening in front of a camera and use that understanding to control a computer or robotic system. Application areas for computer-vision technology include video surveillance, biometrics, automotive, photography, movie production, Web search, medicine, augmented reality gaming, new user interfaces, and many more.

As mobile devices such as smartphones and tablets come equipped with cameras and more computing power, the demand for computer-vision applications is increasing. These devices have become smart enough to merge several photographs into a high-resolution panorama, or to read a QR code, recognize it, and retrieve information about a product from the Internet. It will not be long before mobile computer-vision technology becomes as ubiquitous as touch interfaces. Despite the recent success, there is still significant progress required before we have robots assisting the elderly, cleaning our homes, or fetching household items. A particular challenge for mobile robots in an indoor environment is that most of the objects to be manipulated occupy small portions of cluttered scenes [1].

Consumer-grade range cameras such as the Kinect sensor [2] have the potential to be used in mapping applications where accuracy requirements are less strict. To realize this potential insight into the geometric quality of the data acquired by the sensor is essential.

Low-cost range sensors are an attractive alternative to expensive laser scanners in application areas such as indoor mapping, surveillance, robotics and forensics. A recent development in consumer-grade range sensing technology is Microsoft's Kinect sensor and Intel RealSense. Kinect was primarily designed for natural interaction in a computer game environment. However, the characteristics of the data captured by Kinect [3] have attracted the attention of researchers from other fields including mapping and 3D modeling.

Intel RealSense technology supports a wide range of operating systems and programming languages. The Intel RealSense SDK [4] enables you to extract depth data from the camera and use the interpretation of this data in the platform of your choice. The kit will also offer sample codes, debug tools, and evaluation tools to accelerate your project.

3D modeling (or three-dimensional modeling) is the process of developing a mathematical representation of any surface of an object in three dimensions via specialized software. The product is called a 3D model. It can be displayed as a two-dimensional image through a process called 3D rendering or used in a computer simulation of physical phenomena. The model can also be physically created using 3D printing devices [5]. Three-dimensional (3D) models represent a physical body using a collection of points in 3D space, connected by various geometric entities such as triangles, lines, curved surfaces, etc. [6]. Being a collection of data (points and other information), 3D models can be created by hand, algorithmically (procedural modeling), or scanned [7]. Their surfaces may be further defined with texture mapping.

The focus in this paper is on fusing 2D image information and depth information from stereo images into one model for localization, particularly in the case of contourbased objects. A primary motivation for our work stems from our recent experiences in designing a vision system for our robot ROBKO 18.

2. 3D sensors and software

2.1. Kinect

2.1.1. Characteristics

Even though Kinect for Windows v2 sensor relies on a different technology than Kinect v1, it also allows the acquisition of three different output streams. It is composed of two cameras, namely a RGB and an infrared (IR) camera, from which just one can be seen on the left side of the sensor. Three IR projectors insure the active illumination of the observed scene. Kinect v2 sensor on a photographic tripod. The RGB camera captures color information with a resolution of 1920x1080 pixels, whereas the IR camera is used for the real-time acquisition of depth maps and IR data with a 512x424 pixels resolution. The whole acquisitions can be carried out with a framerate up to 30 Hz. A last feature to be mentioned is the field of view for depth sensing of 70 degrees horizontally and 60 degrees vertically. The technical specifications provided by Microsoft announce an operative measurement range from 0,5 m to 4,5 m.

2.1.2. Specifications

The Kinect sensor consists of an infrared laser emitter, an infrared camera and an RGB camera. The inventors describe the measurement of depth as a triangulation process. The laser source emits a single beam, which is split into multiple beams by a diffraction grating to create a constant pattern of speckles projected onto the scene. This pattern is captured by the infrared camera and is correlated against a reference pattern. The reference pattern is obtained by capturing a plane at a known distance from the sensor, and is stored in the memory of the sensor. When a speckle is projected on an object whose distance to the sensor is smaller or larger than that of the reference plane the position of the speckle in the infrared image will be shifted in the direction of the baseline between the laser projector and the perspective center of the infrared camera.

These shifts are measured for all speckles by a simple image correlation procedure, which yields a disparity image. For each pixel, the distance to the sensor can then be retrieved from the corresponding disparity.

2.2. Intel RealSense

2.2.1. Characteristics

There have been multiple stereoscopic depth cameras released by Intel. However, many of them share similar (or identical) imagers, projectors and imaging processor. We will refer to all these products as the Intel R200, although this analysis also applies to the LR200 and ZR300. These modules are connected and powered by a single

USB connector, are approximately 100x10x4mm in size, and support a range of resolutions and frame-rates. Each unit is individually calibrated in the factory to a subpixel accurate camera model for all three lenses on the board. Distortion and rectification is done in hardware for the left-right pair of imagers, and is done on the host for the color camera.

2.2.2. Specifications

The R200 includes an imaging processor, sometimes referred to as the DS4, which follows a long lineage of hardware accelerated stereo correlation engines, from FPGA systems to ASICs. The image processor on the R200 has a fixed disparity search range (64 disparities), hardware rectification for the left-right stereo pair, and up to 5 bits of subpixel precision. Firmware controls auto-exposure, USB control logic, and properties of the stereo correlation engine.

2.3. SKanect

Skanect is software for capturing a full color 3D model of an object, a person or a room. Skanect transforms your Structure Sensor, Microsoft Kinect or Asus Xtion camera into a low cost 3D scanner able to create 3D meshes out of real scenes in a few minutes. Skanect can acquire dense 3D information about a scene at up to 30 frames per second [8]. Skanect makes it easy to 3D scan different kind of scenes by providing a set of predefined scenarios, suitable for most use cases. Skanect is not just a 3D scanning application. It offers a complete workflow from capture to export and has quite a complete suite of mesh editing tools. The workflow steps are organized in different tabs on the top of the interface.

2.4. ReconstractMe

ReconstructMe is a powerful 3D real-time scanning system. ReconstructMe’s usage concept is similar to that of an ordinary video camera – simply move around the object to be modelled in 3D. Scanning with ReconstructMe scales from smaller objects such as human faces up to entire rooms and runs on commodity computer hardware. ReconstructMe is capable of capturing and processing the color information of the object being scanned, as long as the sensor provides the necessary color stream. ReconstructMe supports a wide range of commodity RGBD sensors such as the ASUS Xtion Family, the PrimeSense Carmine Family or the Microsoft Kinect family. ReconstructMe performs the entire reconstruction in metric space. No need for freaky scaling attempts. The result can be exported to various CAD formats such as STL, OBJ, 3DS, and PLY.

3. 3D modeling method

The three-dimensional modeling method must perform the task of capturing the desired object from different viewpoints until a completed pattern of the object is obtained. The three-dimensional scanning applications work on the principle of recording the measured distances from the sensor to the subject. This is also the reason for using devices that have a deep sensor. An important aspect of the capture process is the resolution of the depth sensor or, more precisely, the number of points to which it measures the distance to the subject. Another very important criterion is to use a powerful video controller that can process data quickly enough and record it. During scanning, applications show frames per second for the capture. The more these frames are, the better and faster the capture is. This also leads to better results in forming the finished three-dimensional model.

The purpose of this study is to apply automated scanning of 3D objects using a mobile robot and a robotic manipulator. Typically, this can also be done manually by a person who holds the scanner and walks around the subject. The experiment will show that there is a difference between the two types of scanning and will explain their advantages and disadvantages.

3.1. Hand manual scan - description

Manual capture is done by man. He must hold the 3D sensor in his hand and walk around the object. Important in the process is that smooth movements must be applied without any sharp deviations. A problem arises when it is necessary to hold or move the sensor cables. We also have a cable length limitation if we use a desktop computer. When crawling all the necessary points of view, the process may be discontinued and the next step of generating the 3D model will be passed.

3.2. Robot manual scan - description

Robot scans are planned to be control by an operator. He will control the robot, which is equipped with its own computer and 3D scanner. The operator can control both the movement of the mobile platform and the movement of the manipulator. This allows smooth movements and quality capture. The robot can be control by a joystick or web based user interface, depending on the environment in which it is working.

4. Results

Following the experiments described in the above paragraph, we have the following results shown in Figures 1, 2, 3 and 4.

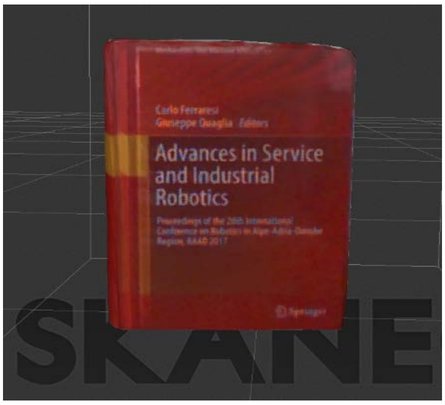

Fig.1 Hand manual scan *Fig.2 Hand manual scan*


Fig. 3 Robot manual scan *Fig.4 Robot manual scan*

The resulting manual scanning patterns have significantly more drawbacks and omissions than those of the robot scanning. The following table shows the advantages and disadvantages of both types of scanning.

1. Conclusion

As can be seen from Table 1, both methods have their advantages and disadvantages. Some of these may be compensated or eliminated if various aids are added. For example, a handheld scanner can use a stand on which to place a 3D scanner. A handler with a high number of degrees of freedom can be used to improve robot scanning. A great advantage of robot scanning is the ability to remotely scan. This is very helpful when working in unfavorable or dangerous rooms.

Table 1. Advantages and disadvantages of scan types

	advantages	disadvantages
hand scan	angle of scanning larger scan range	flicker human error lack of repeatability there is no possibility of scanning remotely
robot scan	repeatability accuracy sustainability keeping a desired position for a long time remote scanning	angle of scanning smaller work area

For the future, we have plans to realize and scan objects with a drone. This will further expand the opportunities for creating 3D objects and contribute to the modeling of large and difficult-to-access outdoor objects. Another plan for future development is to create an algorithm and methods for automated or autonomous scanning that will allow for more accurate 3D models.

Acknowledgments

This research was carry out as part of the project "Telecontrolled Service Robots for Increasing the Quality of Life of Elderly and Disabled, № DN 07/23 – 15.12.2016", financed by the Bulgarian National Science Fund.

References

[1] QiangWu,HuaXu,XukaiZou, ''An effective method for 3D geological modeling with multi-source data integration" Computers & Geosciences Volume 31, Issue 1, February 2005, Pages 35-43.

[2] Wolfgang BOEHLER, Andreas MARBS, ''3D SCANNING INSTRUMENTS" The potential of non-contact close range laser scanners for cultural heritage recording, XVIII. CIPA Symposium, Potsdam, Germany.

[3] Xi Chen, Yuheng Huang, Sida Wang, "3D Scanner using Kinect" Final design report for ESC471 April 25, 2010.

[4] Monica Carfagni , Rocco Furferi, Lapo Governi,Michaela Servi, Francesca Uccheddu,Yary Volpe, "On the Performance of the Intel SR300 Depth Camera: Metrological and Critical Characterization" IEEE Sensors Journal Volume: 17 , Issue: 14 , July15, 15 2017.

[5] FrédéricBosché, "Automated recognition of 3D CAD model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction" Advanced Engineering Informatics Volume 24, Issue 1, January 2010, Pages 107-118.

[6] Aleksey Golovinskiy ,Vladimir G. Kim,Thomas Funkhouser , "Shape-based recognition of 3D point clouds in urban environments" 2009 IEEE 12th International Conference on Computer Vision 29 Sept.-2 Oct. 2009.

[7] V. Blanz, B. Schölkopf, H. Bülthoff, C. Burges, V. Vapnik, T. Vetter. "Comparison of view-based object recognition algorithms using realistic 3D models" Lecture Notes in Computer Science book series (LNCS, volume 1112).

[8] Charles Nock , Olivier Taugourdeau , Sylvain Delagrange, Christian Messier " Assessing the Potential of Low-Cost 3D Cameras for the Rapid Measurement of Plant Woody Structure" Sensors, Volume 13, Issue 12 (December 2013), Pages 15863-17533.

3D-моделирование для распознавания объектов с помощью датчиков глубины

Денис Чикуртев[1, 2], Ивайло Рангелов[2], Найден Шиваров[2], Димитър Карастоянов[1]

[1]*Institute of Information and Communication Technologies – BAS*

[2]*European Polytechnical University - Pernik*

Email: denis@iinf.bas.bg,
rangelov.ivaylo@yahoo.com, **nshivarov@code.bg**, **dkarast@iinf.bas.bg**

Аннотация: *В статье представленные методы трехмерного моделирования объектов. Эти методы используются для распознавания объектов в области сервисной робототехники. Представленные методы - метод ручного сканирования и метод сканирование роботом. В статье описаны 3D-датчики - Kinect и Intel RealSense и 3D программное обеспечение для сканирования и обработки. Для проведения экспериментов по объективному сканированию применяются методы с использованием мобильного робота и Microsoft Kinect для Windows. Результаты показывают преимущества и недостатки методов.*

Ключевые слова: *3D-датчик, 3D-сканирование, распознавание объектов, Kinect, Intel RealSense, мобильный робот*