3D OBJECT RECOGNITION OF CAR IMAGE DETECTION

Rahul Tadake[1], Sumit Chavane[2], Riya Gangurde[3] , Neha Yeola[4]

[1,2,3,4]*Student, BE Electronics and Telecommunication, JSPM's ICOER, Pune, Maharashtra, India.*

Abstract - This project describes the results of experiments on detection and recognition of 3D objects in RGB-D images provided by the Microsoft Kinect sensor. While the studies focus on single image use, sequences of frames are also considered and evaluated. Observed objects are categorized based on both geometrical and visual cues, but the emphasis is laid on the performance of the point cloud matching method. To this end, a rarely used approach consisting of independent VFH and CRH descriptors matching, followed by ICP and HV algorithms from the Point Cloud Library is applied. Successfully recognized objects are then subjected to a classical 2D analysis based on color histogram comparison exclusively with objects in the same geometrical category. The proposed two-stage approach allows to distinguish objects of similar geometry and different visual appearance, like soda cans of various brands.

Key Words: 3D object detection and recognition, Kinect, point cloud analysis, RGB-D images, VFH, CRH, ICP

1. INTRODUCTION

The understanding of the observed environment based on computer registered images and, in particular, finding the number, the type, the properties and finally the pose of objects within this environment is one of the most profound problems and goals that face the machine vision community. Whereas the analysis and interpretation of images and extraction of key information contained therein are most often intuitive, effortless and instantaneous for humans, it is one of the crucial competencies that computer systems still algorithms in this field are in their early infancy due to the enormous complexity of the process and superficial knowledge of its progress in the human brain.

One of the key issues associated with the manipulation of objects is their detection, recognition and localization in the visual scene. The latter task seems to be particularly difficult, however, it became solvable in nearly real time with the application of depth images provided by sensors like the Microsoft Kinect. The Kinect-generated RGB-D image does not only contain the usual three color components of the observed scene for each pixel, but it also holds the distances of the observed points from the sensor. This opens up a whole new range of possibilities for analysis and processing of information, but at the same time, it creates new challenges that require new solutions.

2. LITERATURE SURVEY

1) Abadi M., et al., Tensorflow: Large-scale machine learning on heterogeneous systems, *Software available from tensorflow. org*, 2015.

(2)Aldoma A. et al., CAD-model recognition and 6DOF pose estimation using 3D cues, *IEEE International Conference on Computer Vision Workshops*, 6-13 November 2011.

(3) Aldoma A., et al., A global hypotheses verification method for 3D object recognition, *Computer Vision–ECCV*, 511–524, 2012.

(4) Aldoma A., et al., Multimodal cue integration through Hypotheses Verification for RGB-D object recognition and 6DoF pose estimation, *IEEE International Conference on Robotics and Automation (ICRA)*, 2104–2111, 2013.

(5) Aldoma A., et al., OUR-CVFH – Oriented, Unique and Repeatable Clustered Viewpoint Feature Histogram for Object Recognition and 6DOF Pose Estimation, *Pattern Recognition. DAGM/OAGM 2012. Lecture Notes in Computer Science, vol. 7476*, 113–122, 2012.

4. BLOCK DIAGRAM

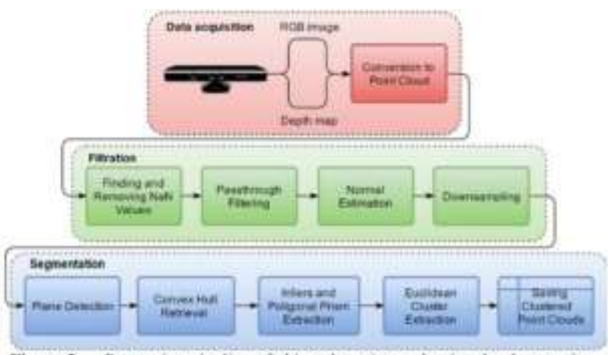

Figure 2. Processing pipeline of object detection and point cloud extraction

5. PROPOSED SYSTEM

A point cloud of the entire scene, generated by the Kinect sensor, has to be preprocessed in such a way that only relevant points belonging to unique objects of interest are grouped together, and stored for later processing. It is relatively easy to detect and remove certain known structures like tables or the floor or any other planar surface of a major size. In this way, all the remaining points can be clustered into separate objects. The pipeline of this process is shown in and processing results at each stage are depicted. A 640 by 480 pixels RGB-D image acquired by the Kinect sensor is processed with the PCL library and converted into a point cloud containing exactly 307 200 point). It requires initial filtration before it can undergo the segmentation process. At first, points that are of no use for 3D processing as having no information about depth (NaN) due to occlusions, transparent or specular surface etc. are removed. Subsequently, a passthrough filter is used to remove all the points lying outside of the user-defined range. Experiments have shown that reliable recognition of small objects is not possible beyond 1.5 m from the sensor and, thus, the passtrough filter cut-off distance along the Z axis was set to this value. For the remaining points in the cloud, normal vectors are computed. Despite the removal of some of the points, the point count in the cloud is still high, which may slow down further processing. Optionally, the point cloud may be downsampled using a voxelized grid method in order to increase performance. Nevertheless, downsampling may have a negative impact on recognition quality, especially in the case of small or distant objects.

6. FUTURE SCOPE

1.In light of the above, it seems necessary to use 2D images with higher resolution, which is possible with the Kinect sensor (maximum image resolution is 1280x1024 RGB) or an optional high-resolution cameras.

2.This would allow to apply more sophisticated 2D image comparison techniques like keypoint detectors and descriptors for image matching.

3.Different objects require different numbers and kinds of models in order to provide the same level of recognition. Having more models does not always result in an increase in the overall system quality.

7. APPLICATION

1. Height estimation the first example is an algorithm that identifies objects and estimates their height off the ground.

2. This is primarily aimed at measuring a person's heightCan be used in Homes.

3. Another example that highlights the advantages of range imaging is a gesture control application. In this application, the user can activate and control a menu system, press virtual buttons, and move virtual slider controls simply by moving an empty hand in free space.

8. CONCLUSION

In this paper, the results of experiments on detection and recognition of three-dimensional objects in RGB-D images provided by the Microsoft Kinect sensor were described. Although the focus was put on using a single image for that purpose, utilizing a series of frames was also considered and evaluated. Experiments performed on hundreds of test scenes show that the proposed and rarely used approach based on the global VFH and CRH descriptors combined with the ICP method and final hypotheses verification can be successfully applied to recognition and localization of objects. However, in order to achieve high recognition rates, the model dataset must be optimized and the distance between the Kinect and the objects being recognized should be relatively small (preferably less than 1 m).

9. REFERENCES

[1] Muhammad Izhar Ramli, Mohd Helmy Abd Wahab, Nabihah, "TOWARDS SMART HOME: CONTROL ELECTRICAL DEVICES ONLINE" ,Nornabihah Ahmad International Conference on Science and Technology: Application in Industry and Education (2006)

[2] N. Sriskanthan and Tan Karand. "Bluetooth Based Home Automation System". Journal of Microprocessors and Microsystems, Vol. 26, pp.281-289, 2002.

[3] E. Yavuz, B. Hasan, I. Serkan and K. Duygu. "Safe and Secure PIC Based Remote Control Application for Intelligent Home". International Journal of Computer Science and Network Security, Vol. 7, No. 5, May 2007.

[4] Pradeep.G, B.Santhi Chandra, M.Venkateswarao, "Ad-Hoc Low Powered 802.15.1 Protocol Based Automation System for Residence using Mobile Devices", Dept.of ECE, K L University, Vijayawada, Andhra Pradesh, India IJCST Vo l. 2, SP 1, December 2011

[5] Sushant Kumar and S.S. Solanki, "Voice and Touch Control Automation", 3rd Int'l Conf. on Recent Advances in Information Technology, 2016