Tracking Objects by a Quadcopter system

Author : Siva Ariram

Guided by : Zdzisław Kowalczuk

Gdansk University of Technology

GDAŃSK UNIVERSITY OF TECHNOLOGY

ABSTRACT

The main goal of the work is to create software for the microcontroller controlling the quadcopter aimed at determining the trajectory of the flight, following the moving object. This is a classic issue of tracking an object through a vision system. Recently, we've got witnessed plenty of industrial momentum in commercializing computer game and increased reality. Head mounted devices (HMD) like Microsoft HoloLens are expected to impacts in diversion, design and medical treatment. whereas plenty of R&D focus is on low-level laptop vision (e.g. localization, reconstruction) and optics design, we tend to believe for powerful future VR and AR applications, linguistics understanding of 3D surroundings can play a very necessary role.

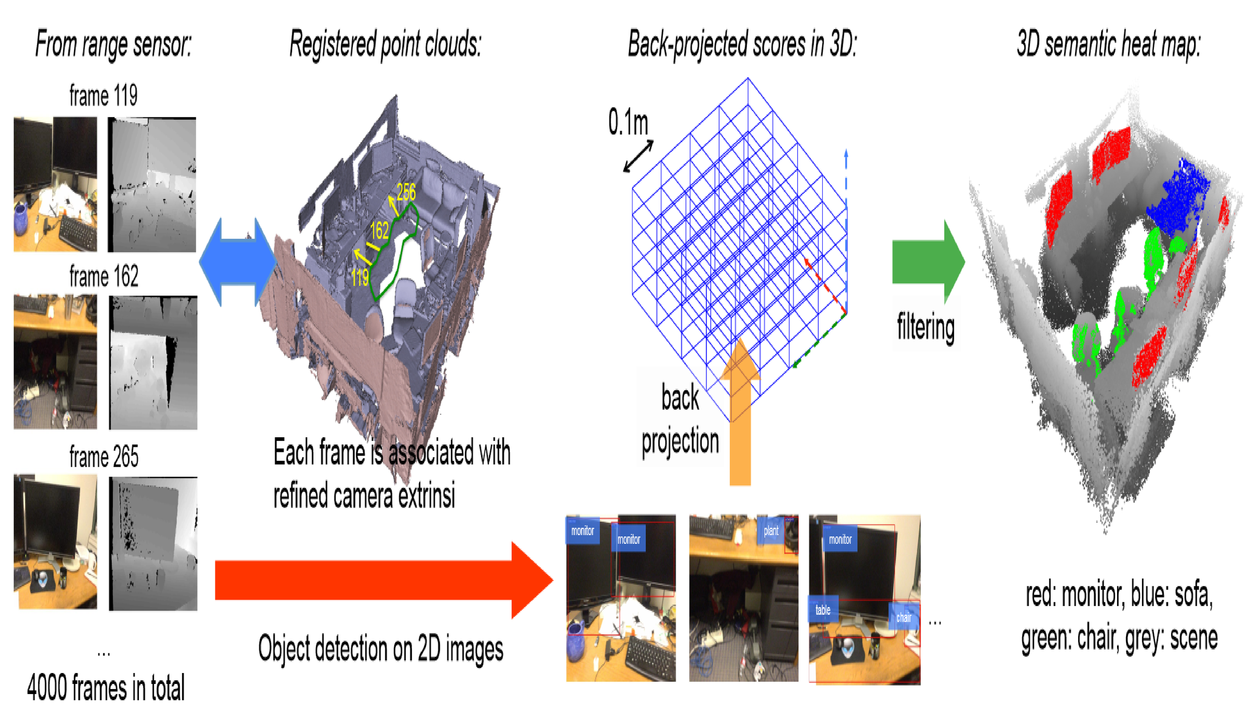
INTRODUCTION

Without linguistics understanding, the pc generated world is truly become independent from the important world. To support interaction of the objects/agents from both worlds, laptop has to assign "meanings" to the pixels and points. for instance, to form a ball rolling in a area, we've got to grasp wherever the ground is. If we would like to form a virtual character sitting beside America on a settee, we've got to grasp wherever the couch is within the area and that a part of it's for sitting. If we would like to create a horror flick in AR to own a ghost setting out from a screen, the algorithmic rule ought to puzzle out whether or not there are any monitors within the area. All of those applications need linguistics understanding.

It's simple for the virtual world since everything is generated and everyone its properties (like class, shape, material, etc.) are identified, whereas it is terribly difficult for the important world. This project can concentrate on a way to provide meanings to real information, i.e. to fill the linguistics gap.

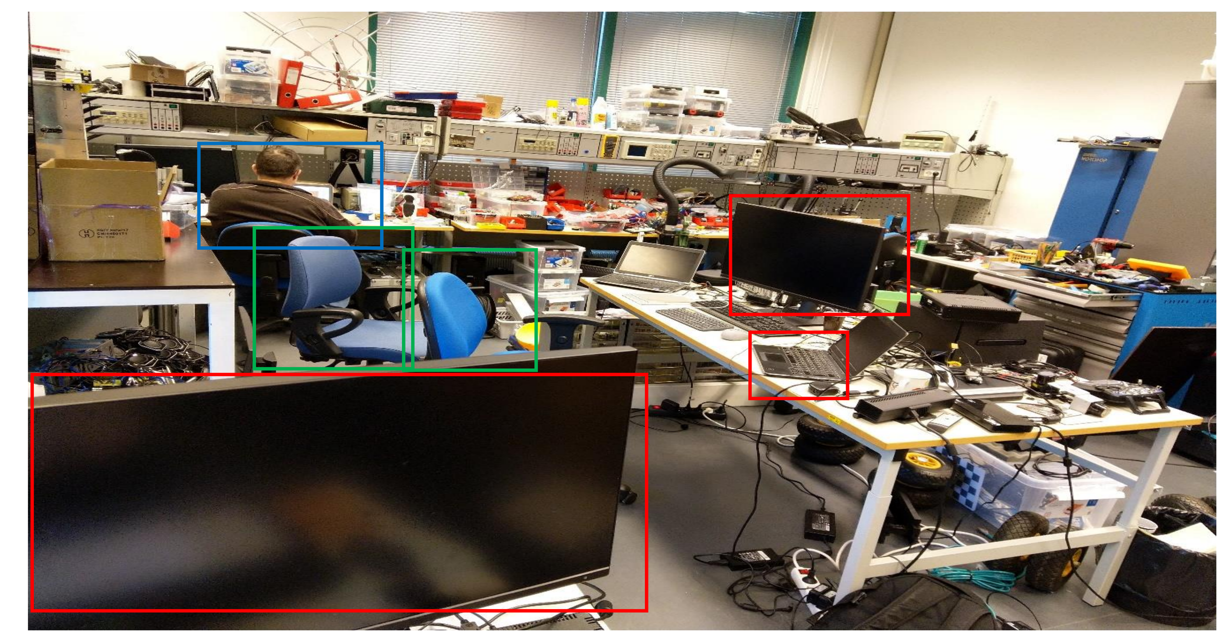
While ancient object detection algorithms are obtainable for RGB pictures, they're not strong enough and can't directly be applied to 3D cases. On the opposite facet, thought seeing ways on purpose clouds don't seem to be data-driven. Usually, key point detection (using handcrafted features), RANSAC and ICP ways are used and that they believe heavily on the idea that a awfully similar model or purpose cloud to the article in real scene is out there, that is typically not the case for generic scenes. Our methodology can avoid drawbacks of those 2 approaches however take their strengths. We have to approach that takes advances in CNNs for perception RGB pictures and 3D as a data aggregation area. By relating object detection on RGB pictures, back project detection scores to post-filtering and world adjustment, we tend to are able to attain strong object detection in 3D scenes. We tend to present the details of the algorithmic rule and we tend to show output results of every step of the pipeline.

METHODOLOGY

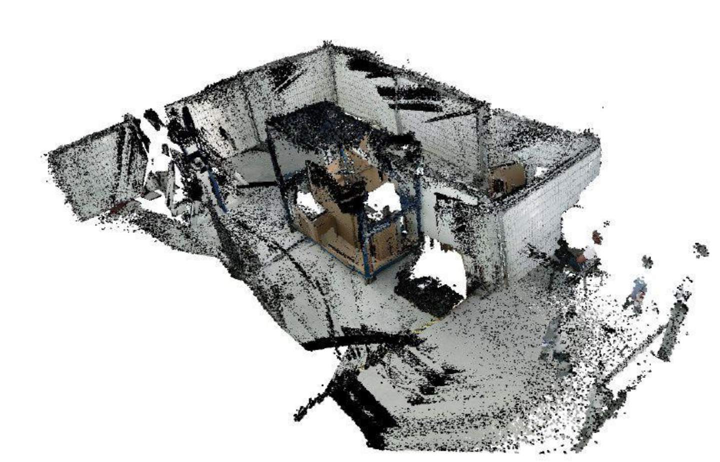


Our methodology first uses the progressive object detector on RGB pictures to attain bounding box proposals beside their scores (note that there are unit heaps of false positives and false negatives during this stage). Then since we've got camera pose and depth image for every frame, we will back project the proposal into 3D voxel grid (from image coordinate to camera coordinate so to world coordinate). All points within the bounding boxes at intervals the depth image area unit projected to their corresponding voxel grid, wherever data is aggregate. Then we tend to conduct applied math filtering and international adjustment (by excluding object collisions) to attain a clean and final heat map.

Object Detection

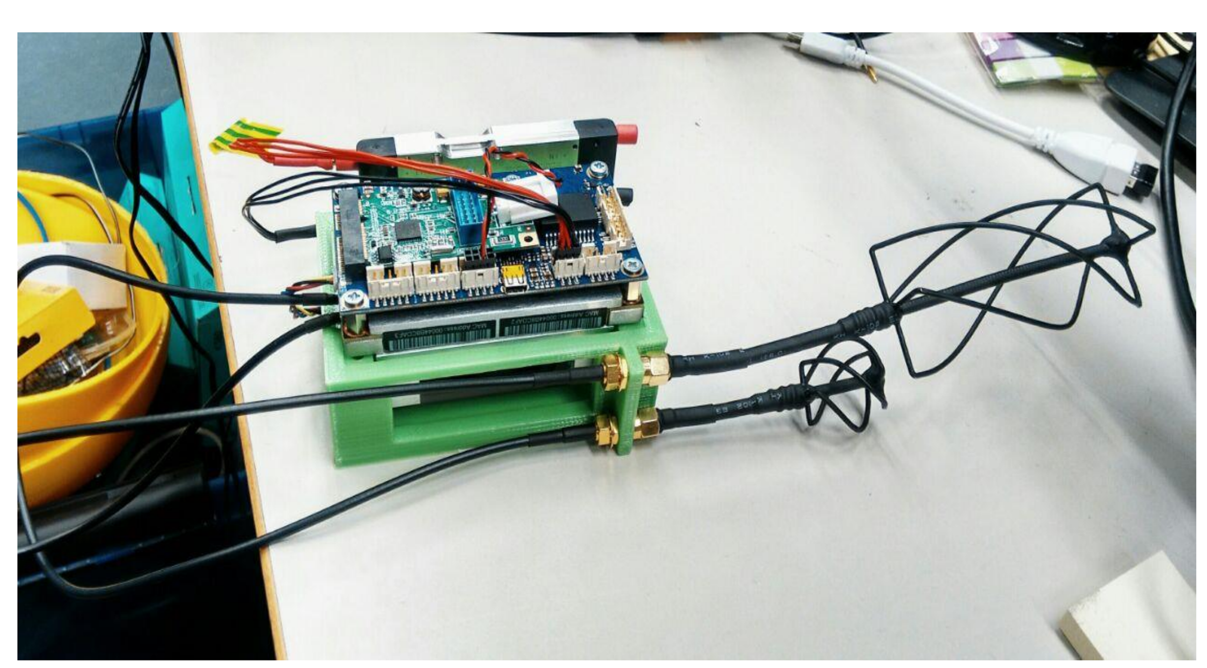


3D Reconstruction

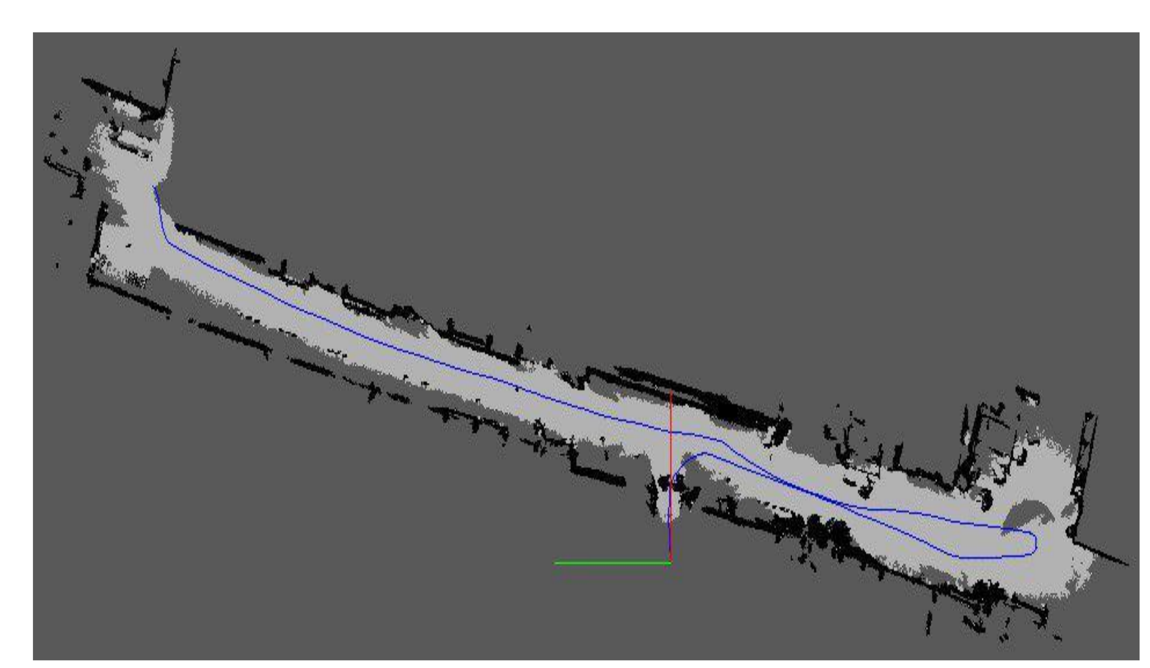


RESULTS

Hardware Implementation



Positional Tracking



CONCLUSION

To sum up, the project has built a quad-copter in the "x" frame configuration and a navigation module and a steering module implemented in it. The vehicle is able to perform basic maneuvers such as elevation, descent, forward tilt, side swing and rotation around the central vertical axis. The project also constructed an operator's tool consisting of a set of microcontrollers and a remote equipped with appropriate sensors and buttons. The device performs both diagnostic and control functions. The selected remote module has a touch display displaying the current parameters such as: vehicle condition, information on orientation and geographic coordinates. When it comes to precision, the results were promising. We showed how the system trained in general image data can be used to detect objects in a specific task (motion detection), thus showing the ability to adapt methods. In many cases, it has detected more objects than annotators the original data has been marked. They were nevertheless marked as false positive clear possession of the appropriate class of objects through visual inspection. On the other hand, the system also made a few mistakes. Water poles, trees and traffic cones were confused with people. We deduced it this was most likely caused by the lack of negative training examples. Come on, we showed the importance of further processing of results. Not maximum damping has a significant effect on medium precision. Generally, overlapping objects and envelopes cause problems with detecting objects systems. We have also shown that the choice of the method of generating the region has an impact both the speed and accuracy of the object detection system.

KEY REFERENCES

1. Henry, P., Krainin, M., Herbst, E., Ren, X., & Fox, D. (2012). RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments. *The International Journal of Robotics Research*, *31*(5), 647-663.
2. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

ACKNOWLEDGEMENTS

Sincere Thanks to Supervisor Prof. Zdzisław Kowalczuk

Special Thanks to Prof. Marek Tatara