SLAM Modelling and Object Detection for Mobile Robot with HoloLens

Reza Moezzi <Reza.Moezzi@tul.cz>, Jaroslav Hlava <Jaroslav.Hlava@tul.cz>

This project shows an innovative application of augmented reality device, Microsoft HoloLens in Simultaneous Localization and Mapping (SLAM) based on computer vision deep learning algorithm for a developed object recognition and obstacle detection task. The most important objective of this implementation is to enhance SLAM modelling by Microsoft HoloLens and detect objects and obstacles through deep learning algorithm named YOLO (You Only Look Once). This system helps the robot to use HoloLens to detect and recognize the obstacles in indoor maneuver. For the object detection, the single forward propagation convolutional neural networks (YOLO) algorithm with three different versions are performed and compared.

Keywords: SLAM, Deep Learning, YOLO, Mobile Robot

Introduction

To navigate an autonomous robot in an unknown environment, localization is necessary to know where the robot is. Localization is the estimation of robot position based on a definite map of environment. In reverse, mapping is the structure of environment's map, knowing the right pathway of the robot [1]. in an indefinite environment for autonomous robot, SLAM (Simultaneous Localization and Mapping) model is essential for navigation [2]. On the other hand, Object detection in machine vision is the process to obtain information, identify and detect objects by computer in the camera's field of view (FOV). In addition, the extracted information must be well organized and easy to use for desired application. The benefits to detect object using video observation systems is already showed its effectiveness as regards the deep learning algorithms development. These systems require more than just a device to recognize and detect the object. First, device stores the scene of the objects as sequences of images then it applies deep learning algorithm to deliver the result of recognition and to detect the object. Nowadays with advanced technologies for augmented reality (AR) such as Microsoft HoloLens, it can visualize, recognize, and detect, the objects in front of the user or the Robot, in real world. Spatial mapping feature in Microsoft HoloLens delivers a detailed image of real-world surfaces in the environment around it The modern deep learning algorithms such as Convolutional Neural Network (CNN) [3], Region-based CNN (R-CNN) [4], fast R-CNN [5], and You Only Look Once (YOLO) [6] can be implemented in Microsoft HoloLens for object and obstacle detection, and it enables the robot to identify objects and obstacles in AR area.

Methodology

In this section, the most important methodology and procedure of the system is described to make sure proper object and obstacle detection via HoloLens with YOLO. As shown in Fig.1, the software architecture is based on server-client connection. The hardware used are, two parallel NVIDIA GPU (graphic processing Units) Quadro P-4000 to run on the server which process the YOLO algorithm. HoloLens v1 and its built-on cameras are used as an input scene for object detection in the environment in the client side.

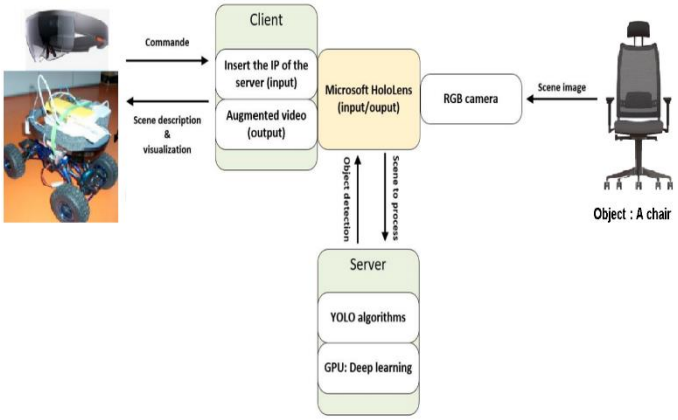

Fig 1: Architecture of the system

Result and Discussion

Our described approach to recognize and to detect the object and obstacles via HoloLens indicate an exceptional detection result. YOLO v1, YOLO v2 and YOLO v3 are used to obtain the results in term of processing time and precision mean average. For higher efficiency in detection procedure, for each mentioned version of YOLO, we have defined specific fps (number of frames per second). This method's performance is based on finest precision with the finest time to recognize and detect the object. For the measuring precision, YOLO with three versions is used to confirm which one is the most accurate one. As shown in Fig.2, The findings show the effectiveness of YOLO for the precision average of various objects in our workplace. Based on the results, YOLO v3 gave a good precision detection which is higher than 90%.

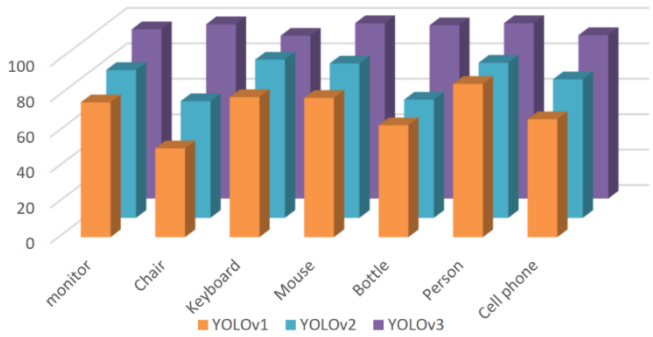

Fig 2: Percentage of precision with respect to different detected objects by different YOLO versions

Conclusion

The SLAM is an essential tool in the field of autonomous mobile robot while it navigates desired path, simultaneously creating a map of the physical environment and the vehicle estimates position of itself on corresponding map. This paper illustrated the use of computer vision deep l learning algorithm called YOLO to apply on HoloLens for an enhanced SLAM. we have designed a system between desktop computer as a server and HoloLens device attached to the mobile robot, as a client side which are connected via TCP/IP channel. The system aimed to process the object and obstacle detection with HoloLens utilizing a server based on YOLO in a real environment. The results present an effective and accurate detection by HoloLens with 96% mean average precision and 5 fps as the detection processing time, using YOLO v3.

Acknowledgments

This work was supported by the Student Grant Competition (SGS) project at the Technical University of Liberec in 2021.

Reference

[1] R. Siegwart, I. Nourbakhsh and D. Scaramuzza, Introduction to Autonomous Mobile Robots, London, England: The MIT Press, 2011.

[2] J. Sasiadek, A. Monjazeb and D. Necsulescu, "Navigation of an autonomous mobile robot using ekf-slam and fastslam," 16th Mediterranean conference on Control and automation, p. 517–522, 2008.

[3] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2323. doi:10.1109/5.726791

[4] Guo, J., Xu, J., Liu, S., Huang, D., & Wang, Y. (2016). Occlusion-robust face detection using shallow and deep proposal based faster R-CNN doi:10.1007/978-3-319-46654-5_1

[5] Girshick, R. (2015). Fast R-CNN. Paper presented at the Proceedings of the IEEE International Conference on Computer Vision, , 2015 International Conference on Computer Vision, ICCV 2015 1440-1448. doi:10.1109/ICCV.2015.169

[6] Redmon, J., Santosh D., Ross G., and Ali F.: You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788 (2016).