Kinect Depth sensor for Computer Vision Applications in autonomous Vehicles

Niyonsaba Eric
Department of Computer Engineering
Dong-Eui University
Busan, Korea
E-mail:niyonsabaeric@gmail.com

Jong-Wook Jang
Department of Computer Engineering
Dong-Eui University
Busan, Korea
E-mail:jwjang@deu.ac.kr

***Abstract*— Unmanned vehicles (Smart cars, drones, robots) need to understand and respond to the surrounding environment in order to perform their tasks. Therefore, they must be equipped with vision capabilities (vision sensors) to let them to detect the presence of objects in their paths or measure how far away objects are from the sensor depending of the type of application. Several technologies have been used to deal with this computer vision application such stereovision, LIDAR and RADAR but each of them still has advantages and disadvantages in term of limitations and price. In this paper, we developed one of the functions of unmanned vehicles to detect the presence of objects and measure how far they are using Kinect depth sensor, a low-cost range sensor along with its higher depth fidelity and attractive alternative in computer vision. The results showed that Kinect sensor can detect along with segmentation techniques the presence of objects in its field of view and measure how far they are from the sensor. Based on results, Kinect can be mounted on unmanned vehicles like a vision sensor for obstacle avoidance application or other applications which require a vision sensor to measure how far detected objects are from it as well for also manned vehicles to alert drivers but mostly computer vision has a crucial role in vehicles without human intervention.**

Keywords*— *Kinect depth sensor, Kinect distance measurement, Vision sensor, depth segmentation technique, object detection.

I. INTRODUCTION

Naturally, vision comes for human beings with their two eyes and brain to understand the surrounding environment and make decision according to what they see. Computer vision works in the same perspective using vision sensors on autonomous systems along with some sophisticated algorithms to extract and analyze useful information from images and video sequences. Computer vision has been an interesting field and is undergoing research so far and its applications are numerous including collision avoidance [1], autonomous vehicles, industry quality inspection, face recognition, medical image analysis, surveillance systems etc. One of the integral tasks in computer vision is to perceive the word as human do, detect objects in vision sensor's field of view and estimate the distance from where those objects are located. Many techniques along with different vision sensors have been used for computer vision applications including Stereovision, LIDAR (Light Detection and Ranging) and RADAR (Radio Detection Ranging) [2] but still each of them has its own advantages and disadvantages. Finding the distance between the camera and detected objects within camera's field of view, which is the main purpose of the present research, is an ongoing field of research given the fact that there isn't a perfect solution yet and computer vision prefers to use depth cameras rather image intensity cameras since depth information makes the variety of applications more feasible and robust [3]. In this paper, we chose to make experiences with Kinect sensor, a low-cost depth sensor for detecting objects in its field of view and measuring how far away they are from the sensor. Computer vision has been used in many specific applications such as multimedia database search engines [4] using attributes such as colors and shapes instead of text-based retrieval queries for images. In medicine, computer vision can alert clinicians and assist doctors during results interpretation to reveal some abnormalities on medical images [5]. Surveillance systems are also based on computer vision techniques in public places like airports and transport stations. In factories, computer vision can even monitor workers and track anyone who is not paying attention to a potentially dangerous part of a task [6].

Depending on these applications mentioned above, comput er vision relies on different vision devices and techniques. In t he literature review, we can so far identify many applications r elated to computer vision with different vision devices. Comp uter vision is improving in the technology with range cameras in human-computer interaction, robotic and machine vision, a utonomous vehicles as well as in augmented reality. Accordin g to Jernej Mrovlje and Damir Vrancic [7], distance measuring sensors are divided into active and passive categories. Active methods (laser beam, radio signals, etc.) used as geometric sen sing, measure the distance by sending some signals to the surf ace of the object whereas passive ones receive information fro m detected object using light. In their experiments, they used a passive technique, stereoscopic measurement method to find distance to object but this method requires two cameras and ha s more restrictions. Giulio Reina and Annalisa Milella [8] hav e experienced making agriculture robots based on multi-baseli ne stereovision. They found out that even the system has adva ntages, there was a need of good light to make clear the field o f view and also its algorithms are very complicated with a hig h computation cost which makes slow systems in real time. Th e background and the obstacle are very hard to separate. LIDA

R and RADAR sensors have been used as obstacle detection for autonomous vehicles but they are very expensive and this prevents researchers to use such sensors.

II. Kinect Architecture and related work

Kinect camera was released by Microsoft Corporation to interact with game consoles at the beginning of 2010 and the major idea behind was to develop an interface device to avoid hand- controller devices using depth data to transform player's movement into controls and also voices for spoken commands. The first version was released with Xbox 360 console. The second version was released in 2014 along with Xbox One console and more performances compared to the first one.

Kinect sensor version 2 appears as a popular sensing input device that serves as a natural user interface application for computers and game consoles (Xbox one). Kinect v2 sensor is used in many different fields of technology as it can sense depth, capture color images, emit infrared rays, and input audio [9].

Fig.1. Kinect version 2 architecture

The color camera is responsible for capturing and streaming the color video in order to detect the red, blue and green colors from Kinect with a resolution of 1920x1080 pixels whereas the depth sensor generates the depth information of the object in front of the Kinect with a resolution of 512x424 pixels [10]. Meanwhile the infrared emitter (IR camera and IR projector) allows the sensor to view light-independently. The new depth sensor for Kinect v2 has a higher depth, fidelity and a significantly improved noise floor, and by providing 3D visualization, we are able to see smaller objects and all objects more clearly [11]. As Kinect v2 uses Time-of-Flight system, it modulates the camera light source with a square wave by using phase detection to measure the time the light takes to travel from the light source to the object and back to the sensor [12]. The depth is then calculated based on the speed of light in air.

Even though Kinect v2 was mainly developed by Microsoft as interface for gaming with Xbox One console, its depth sens-or based on Time-of-Flight principle has made it popular in the scientific community and researchers have tested and used it in many fields. In [13], authors have tested the capacity of Kinect v2 in order to find out if it could be an alternative to laser scanners for 3D measurements. They concluded that achieved results based on measurement precision and outdoor efficiency and that Kinect can be seen as a real progress for computer vision applications. Another test in [10] have been made for investigating Kinect's properties including depth accuracy and authors obtained that Kinect has a good accuracy for objects which are placed near the sensor. Also, authors found that there are some parameters which limit Kinect's performance such as object with reflective material like mirror and light-absorbing materials. Those kind of objects make the IR (Infrared) light emitted by Kinect sensor difficult to be reflected back to the camera for depth measurement.

To sum up, Kinect for windows sensor v2 shows with its low price a great potential to be applied in computer vision applications. Therefore, we can find in the literature review some of researchers using Kinect v2 in different applications. Tracking an object's 3D location has been possible at low-cost by using Kinect sensor[14] whereas 3D object tracking was reserved before for users who can afford high-cost motion tracking systems such as Vicon system which is priced upwards of $10.000. Authors showed that object tracking can be performed in real time by using depth and color data from Kinect sensor priced cheaply at around $129. Tracking object has many applications in computer vision such as surveillance [15], pose and facial recognition [16] etc. and this makes Kinect a useful tool in many fields.

Kinect for windows sensor v2 has been also used for autonomous vehicles such as in SLAM (Simultaneous Localization and Mapping) systems [17] which is a major function in computer vision and robotics by using both depth and color images from a low cost sensor. Also in [6], Kinect has been mounted on a mobile robot for navigation and authors appreciated such system which allows mapping algorithms to model the environment and detect obstacles. In [18], Kinect has been used in healthcare systems for blind navigation support system and real-time monitoring of activities for people with mental illnesses which is a difficult task when it comes to be operated by a clinician. Kinect was used also in agriculture activities [2] to reduce human involvement for harmful activities in a greenhouse environment such as using some products for spraying or fruit collection which requires a high degree of autonomous system. It is obvious that Kinect v2 has attracted many researchers in different domains and seems producing promising results. It seems that lots of application deal with object detection techniques for obstacle avoidance but here, in this paper, another parameter is added, the distance value from the detected object to the sensor for further actions within Kinect's field of view. Different scenarios are shown based on different fields of view obtained by slicing Kinect's field of view and detected object's size and results are discussed in the next section.

III. Results and discussion

The main purpose is to detect objects within Kinect sensor's field of view and take their depths from where they are located. Here, Kinect depth sensor as shows Fig.2 is considered as a vision sensor connected to a computer system for distance measurement function of an entire autonomous system. Design system shows whole field of view of Kinect camera and also target area delimited by depth range to test the consistency of Kinect depth sensor in different intervals.

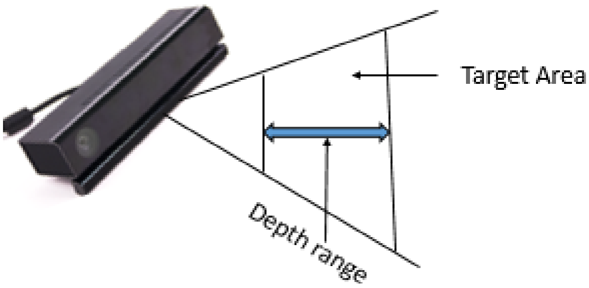

Fig.2 System Design of distance measurement function

As long as frames are streaming from Kinect camera, present objects in its field of view are identified with a red rectangle using image processing functions and the depth sensor processes the distance measurement system from a nearest pixel of each object.

In this section, results from our system using Kinect sensor in front of some objects are shown along with some settings indicating in which range user wants to detect objects and how should be the minimum size of tracked object because size have important effects in terms of processing time specially for real-time applications.

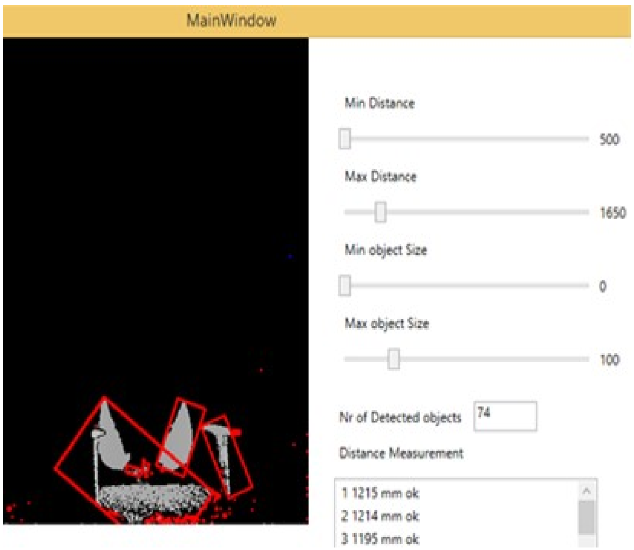

Fig.3. Tracked objects

The first impression here is the number of 74 tracked objects within 500mm to 1650mm range. As show on Fig.3, this is because the minimum size of objects is set to zero and Kinect tends to identify some simple pixels which are not really objects but simply can be considered as noise as we can note on the corresponding color image below on Fig.4.

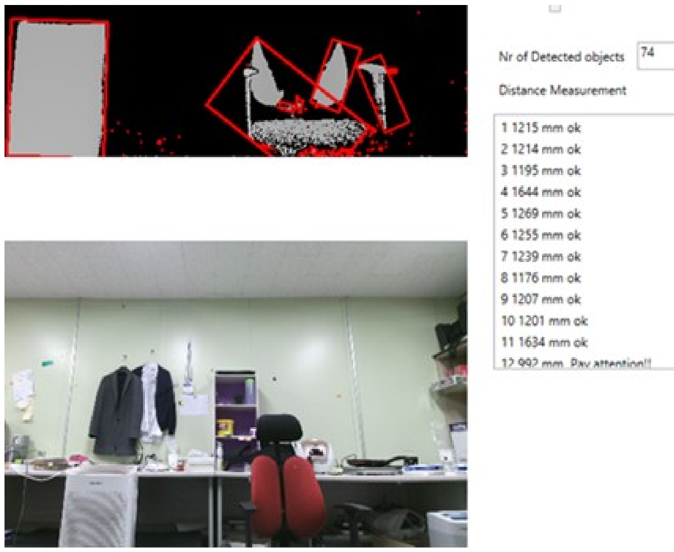

Fig.4 Corresponding color image

The drawback is that Kinect takes much time to process each frame for distance measurement as objects are increasing. To overcome this, the interface of this system let the user to set the reasonable minimum size for detected objects and indicate the range in which Kinect should detect objects instead of the whole field of view.

Therefore, when the settings are modified, the minimum size to 13, we could find for the same scene 4 detected objects which are matched with the real present objects as shows the depth image on Fig.5.

Fig.5. Detected objects with minimum size set to 13

Along with detected objects, also their distances are displayed. Distance is another type of data needed by unmanned vehicles in order to achieve their goals safety. Autonomous vehicles can also need to measure how far are detected objects for making an alert for example or for further processing depending on the type of application.

IV. Depth Accuracy Evaluation

Based on our depth results, there is a variation of depth value as frames are streaming from Kinect sensor. In order to assess the impact for that effect, we recorded depth values for one detected object within 30 first frames in different ranges or intervals. Different intervals have been taken (500mm to 1982mm, 500mm to 3069mm and 500mm to 4500mm) as values are shown below.

These are different distances recorded for one detected object in 30 successive frames from Kinect sensor in different ranges intervals as show tables below.

Table1. Recorded distances in the range from 500mm to 1982mm

1951	1946	1944	1947	1947	1950
1948	1947	1943	1949	1946	1959
1945	1951	1950	1946	1948	1948
1945	1946	1949	1949	1946	1946
1945	1945	1951	1944	1950	1951

Table2. Recorded distances in the range from 500mm to 3069mm

2909	2912	2900	2909	2907	2906
2907	2906	2906	2908	2908	2906
2903	2908	2907	2905	2906	2909
2908	2906	2910	2909	2908	2910
2908	2909	2903	2903	2906	2903

Table3. Recorded distances in the range from 500mm to 4500mm

4404	4408	4408	4410	4407	4402
4408	4410	4407	4407	4406	4407
4403	4406	4414	4405	4404	4409
4408	4413	4411	4406	4409	4410
4405	4403	4404	4427	4409	4404

As those tables show, we can realize that there is variation of depth values for the same one detected object observed from a frame to another. We took different intervals because, in the previous related works, it was found that Kinect's accuracy depends on how far away the target is in Kinect sensor's field of view.

Comparing values obtained for each scenario, it is noticeable that 8mm, 9mm and 25 mm are respectively the maximum variations which give us promising results when using Kinect depth sensor to measure distances for objects located in its field of view. Therefore, one think we should understand is that for applications which are very sensitive to distance measurement or which cannot tolerate at least a variation of 10 mm, Kinect sensor seems to be not useful in that case.

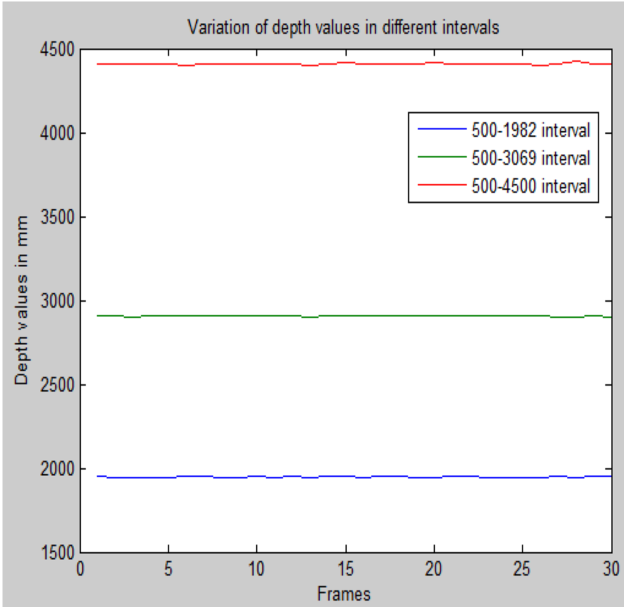

Fig.6. Representation of depth variations in different intervals

As shows Fig.5, even though there is a variation as frames are streamed one after another, Kinect still a useful device for distance measurement within its field of view and we assume that we can get promising results for our next step to mount Kinect sensor on Robot for obstacle avoidance application at a certain distance from the objects in front of it.

V. Conclusion and Future Work

In this paper, by using Kinect depth sensor, we developed one of the needed functions in autonomous vehicles for the vision part. The developed function consists of detecting objects within Kinect's field of view and measuring the distance from Kinect depth sensor to the detected objects. We tried to detect in a very accurate way the presence of objects in front of Kinect sensor taking in account the size of tracked objects to ensure if it is a real object or a pixel noise to reduce the processing time for distance measurement. Also, we tested Kinect depth sensor accuracy and found promising results in its field of view. In the future work, our plan is to use this function of object detection and distance measurement to implement a practical application, a robot equipped with Kinect depth sensor for obstacle avoidance at a certain distance in order to trace the best path for achieving safely its mission.

Acknowledgment

This research was supported by the 2015 Human Resource Development Project for Local Innovation and Creativeness (NRF-2015H1C1A1035898) and the 2016 Human Resource Development Project for New Local Industry

(No.2016H1D5A1910985) of the National Research Foundation of Korea.

References

[1] Stefania Monica, Gianluigi Ferrari, Low-complexity UWB-based collision avoidance system for automated guided vehicles, ICT Express, Vol. 2,no. 2, June 2016, pp. 53-56

[2] Sharon Nissimov, Jacob Goldberger and Victor Alchanatis, "Obstacle detection in a greenhouse environment using the Kinect Sensor". Journal of Computers and Electronics in Agriculture, Vol.113, no.3, Feb. 2015 pp. 104-115.

[3] Burak Gokturk, Hakan Yalcin and Cyrus Bamji, "A Time-Of-Flight Depth Sensor – System Description, Issues and Solutions". Computer Vision and Pattern Recognition Workshop (CVPRW'04), pp.35-44,2004.

[4] Thomas Funkhouser, Patrick Min, Michael Kazhdan, Joyce Chen , Alex Halderman, David Dobkin and David Jacobs, "A Search Engine for 3D Models". Journal of ACM Transactions on Graphics (TOG), Vol.22, no1, Jan. 2003, pp.83-105.

[5] Petrick, Berkman Sahiner, Samuel , Alberto, Loredana, Silvia and Matthew, "Evaluation of computer-aided detection and diagnosis systems". Medical physics Journal, Vol.40, no8, Aug.2013.

[6] Research at the University of Maryland, "Computer Vision: Visual-Event Recognition for security and science". Available online: http://www.research.umd.edu/sites/default/files/documents/brochures/computer-vision.pdf.

[7] Manaf A.Mohamed, Amera I.Melhum and Faris A.Kochery, "Object Distance Measurement by Stereo Vision". International Journal of Science and applied Information Technology, vol.2,no.2, March 2013, pp.5-8.

[8] Giulio Reina and Annalisa Milella, "Towards Autonomous Agriculture: Automatic Ground Detection Using Trinocular Stereovision", Sensors(Basel) journal, vol.12,no.9, Sept. 2012, pp.12405-12423.

[9] Microsoft, Kinect hardware key features and benefits. Available online: https://developer.microsoft.com/en-us/windows/kinect/hardware

[10] Lin Yang, Longyu Zhang, Abdulhameed Alelaiwi and Abdulmotaleb El Saddik.," Evaluating and Improving the Depth Accuracy of Kinect for Windows v2". IEEE Sensor Journal, Vol.15, no.8, Aug. 2015, pp.4275-4285.

[11] ALABBASI, Florica Moldovenau, and Alain Moldoveanu, "Facial Emotion Expressions Recognition with Brain Activities using Kinect sensor v2". International Research Journal of Engineering and Technology (IRJET), Vol.2, no.2, May 2015, pp.421-428.

[12] John Sell and Patick O'Connor, "The Xbox One system on a chip and Kinect sensor". IEEE Micro, Vol.34, no.2, March 2014, pp. 44-53.

[13] Elise Lachat, Hélène Macher, Tania Landes, and Pierre Grussenmeyer, "Assessment and Calibration of a RGB-D Camera (Kinect v2 Sensor) Towards a Potential Use for Close-Range 3D Modeling," Remote Sensing, vol. 7, no. 10, Oct.2015, pp. 13070-13097.

[14] Michael Fleder, Sudeep Pillai and Jeremy Scott, "3D Object Tracking using the kinect", MIT Computer Science and Artificial Intelligence Laboratory. Available Online(Accessed March 2017): https://people.csail.mit.edu/spillai/data/papers/cvclass-project-paper.pdf

[15] Ravishankar Sivalingam, Anoop Cherian,Joshua Fasching, Nicholas Walczak, Nathaniel Bird,Vassilios Morellas, Barbara Murphy, Kathryn Cullen, Kelvin Lim, Guillermo Sapiro and Nikolaos Papanikolopoulos, "A Multi-sensor visual tracking system for behavior monitoring of At-Risk children", IEEE International Conference on Robotics and Automation, pp.1345-1350, 2012.

[16] Karl Pauwels1, Vladimir Ivan, Eduardo Ros1and Sethu Vijayakumar, "Real-time object pose recognition and tracking with an imprecisely calibrated Moving RGB-D camera", IEEE/RSJ International conference on intelligent Robots and systems, pp.2733-2740.

[17] Olivier Wasenmuller, Maecel Meyer and Didier Stricker "CoRBS:Comprehensive RGB-D Benchmark for SLAM using Kinect v2",IEEE Winter Conference on Applications of Computer Vision(WACV), pp.1-7, 2016.

[18] Victor Filipe, Hugo Fernades, Filipe Fernandes and Joao Barroso, "Blind navigation support system based on Microsoft Kinect", Procedia Computer Science , vol.14, pp.94-101, 2012.