JKIICE Journal of the Korea Institute of Information and Communication Engineering

한국정보통신학회논문지(J. Korea Inst. Inf. Commun. Eng.) Vol. 21, No. 6 : 1237~1242 Jun. 2017

키넥트(Kinect) 윈도우 V2를 통한 사물감지 및 거리측정 기능에 관한 연구

니욘사바에릭[1] · 장종욱[2*]

Study on object detection and distance measurement functions with Kinect for windows version 2

Niyonsaba Eric[1] · Jong-Wook Jang[2*]

[1]Department of Computer Engineering, Dong-Eui University, Busan 47340, Korea
[2]Department of Computer Engineering, Dong-Eui University, Busan 47340, Korea

요 약

컴퓨터 비전은 인공 지능 기술을 통해 인간의 시각 시스템을 모방해 주변 환경을 보다 정확하게 인식하는 새로운 이미지 센서 기능으로 각광받고 있다. 본 논문에서는 사물감지 및 거리측정 기능이 있는 새로운 깊이 센서인 키넥트(Kinect) 카메라를 통해, 무인 또는 유인 차량, 로봇 및 드론 등을 위한 컴퓨터 비전의 가장 중요한 기능들을 대상으로 시험을 진행하였다. 키넥트 카메라를 통해 시야 내에 있는 사물의 자리 또는 위치를 예측하고, 실제 사물이 아닌 픽셀을 무시해 처리 시간을 줄일 수 있도록 감지한 사물이 실제 사물인지 확인하여 깊이 센터를 통해 정확하게 거리를 측정한다. 실험 결과, 해당 거리센서는 좋은 결과를 나타냈으며, 추가 프로세싱을 위한 컴퓨터 비전 어플리케이션의 핵심 기능인 사물감지와 거리측정에 키넥트 카메라를 사용한다.

ABSTRACT

Computer vision is coming more interesting with new imaging sensors' new capabilities which enable it to understand more its surrounding environment by imitating human vision system with artificial intelligence techniques. In this paper, we made experiments with Kinect camera, a new depth sensor for object detection and distance measurement functions, most essential functions in computer vision such as for unmanned or manned vehicles, robots, drones, etc. Therefore, Kinect camera is used here to estimate the position or the location of objects in its field of view and measure the distance from them to its depth sensor in an accuracy way by checking whether that the detected object is real object or not to reduce processing time ignoring pixels which are not part of real object. Tests showed promising results with such low-cost range sensor, Kinect camera which can be used for object detection and distance measurement which are fundamental functions in computer vision applications for further processing.

키워드 : 깊이 분할, 키넥트를 통한 거리 측정, 키넥트 깊이 측정 센서, 사물 감지

Key word : Depth segmentation, Distance measurement with kinect, Kinect depth sensor, Object detection

Received 26 May 2017, **Revised** 27 May 2017, **Accepted** 03 June 2017

* **Corresponding Author** Jong-Wook Jang (E-mail:jwjang@deu.ac.kr, Tel:+82-051-890-1709)
Department of Computer Engineering, Dong-Eui University, Busan 47340, Korea

Open Access **https://doi.org/10.6109/jkiice.2017.21.6.1237** print ISSN: 2234-4772 online ISSN: 2288-4165

Ⅰ. INTRODUCTION

Firstly, Kinect camera made by Microsoft corporation was released as an interface device to interact with consoles in order to avoid hand-controller devices during games to transform player's movements and voices into controls with its capabilities to stream depth image, color image, emit infrared and input audio. Therefore, Kinect was basically used for tracking human body's movements, gestures and spoken commands during game but now due its depth sensor, it comes as essential tool in computer vision field. In this paper, using kinect's depth sensor and its high frame rate, we made experiments for vision applications which need to perceive their surroundings. Some other technologies have been used for computer vision applications including stereovision which consist of a set of cameras to estimate the depth of objects in front of it but with some limitations about lack of good light and the complexity of the algorithm. Also, LIDAR(LIght Detection and RAnging) and RADAR(RAdio Detection and RAnging) [1] have been used in high systems because of its accuracy like military services but they are very expensive even more than an entire autonomous vehicle in some cases. Here, I chose to make experiments with kinect camera, a low cost depth sensor, to detect objects within its field of view and measure how far they are which is an ongoing research and computer vision prefers to use depth cameras rather image intensity cameras. Computer vision have been used in many specific domains such as in medicine when it can help for alerting clinicians, assist doctors during results interpretation to reveal some abnormalities on medical images and in surveillance systems in public places like airports and transportation 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 [2].

Depending on these applications mentioned above, computer vision relies on different vision devices and techniques. In the literature review, we can so far identify many applications related to computer vision with different vision devices. Computer vision is improving in the technology with range cameras in human-computer interaction, robotic and machine vision, autonomous vehicles as well as in augmented reality. According to Jernej, Mrovlje and Damir Vrancic [3], distance measuring sensors are divided into active and passive categories. Active methods (laser beam, radio signals, etc.) used as geometric sensing, measure the distance by sending some signals to the surface of the object whereas passive ones receive information from 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 has more restrictions. Giulio Reina and Annalisa Milella [4] have experienced making agriculture robots based on multi-baseline stereovision. They found out that even the system has advantages, there was a need of good light to make clear the field of view and also its algorithms are very complicated with a high computation cost which makes slow systems in real time.

Lots of applications seem to deal with object tracking but in this paper another parameter is added, distance value from detected objects which can be helpful for unmanned vehicles or when they reach at a certain distance from detected objects using kinect depth sensor.

Ⅱ. STUDY SPECIFICATIONS AND RELATED WORKS

Kinect sensor version 2 appears as a popular sensing input device that serves as a natural user interface application for computers and game consoles such as XboxOne and has a higher depth, fidelity and significantly improved noise floor, and by providing 3D visualization to enabling to see objects more cleary [5]. The main purpose is to detect objects within Kinect sensor's field of view and take their depths from where they are located.

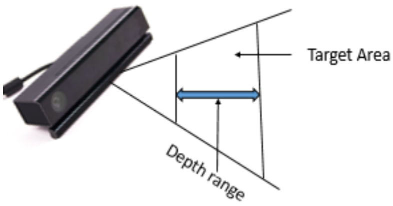

Fig. 1 System Design of distance measurement function

Here, Kinect depth sensor as shows Fig.1 is considered as a vision sensor connected to a computer system for distance measurement function. 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.

Kinect depth sensot is based on Time-of -Flight principle and some researchers have used it in different fields. In [6], kinect has been tested in order to find out if it could be an alternative to laser scanners for 3D measurements and authors concluded that can be seen as a real progress for computer vision applications based on achieved results about measurement precision and outdoor efficiency. 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[7] whereas 3D object tracking was reserved before for users who can afford high-cost motion tracking systems such as Vicon system.

Kinect for windows sensor v2 has been also used for autonomous vehicles such as in SLAM (Simultaneous Localization and Mapping) systems [8] which is a major function in computer vision and robotics by using both depth and color images from a low cost sensor and also in monitoring applications [9]. In [10], kinect camera 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. In this paper, for distance measurement function, different scenarios have been made by slicing kinect's field of view and modifying detected object'size and results are shown in the next section.

Ⅲ. RESULTS AND DISCUSSION

Once depth frame is streamed from kinect depth sensor, our system can detect the present objects identified with a red contour within its field of view and ignore some pixels which are not a part of a real object in order to minimize the processing time. Results 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.

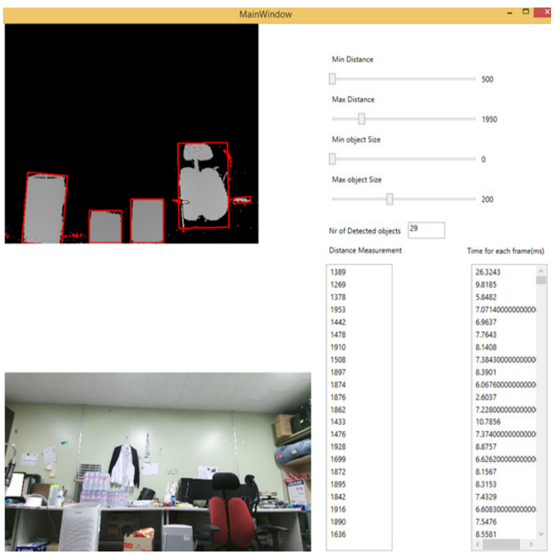

Fig. 2 Tracked objects

The first impression here (shows Fig.2) is the number of 29 tracked objects within 500mm to 1950mm range. 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.

The drawback of this is that Kinect takes much time to process all detected objects but here we need real time system. To overcome this kind of drawback is to set a reasonable minimum size of detected object which can be real object in order to minimize the processing time. The processing time of each depth frame is shown. The first frame takes much time (26 ms) because of “background noise” activity in .NET framework. Therefore, the first time a method is launched in .NET, the time taken by that activity is added to the time of execution.

The minimum size of detected object is increased to 13 and the result can be seen on Fig.3.

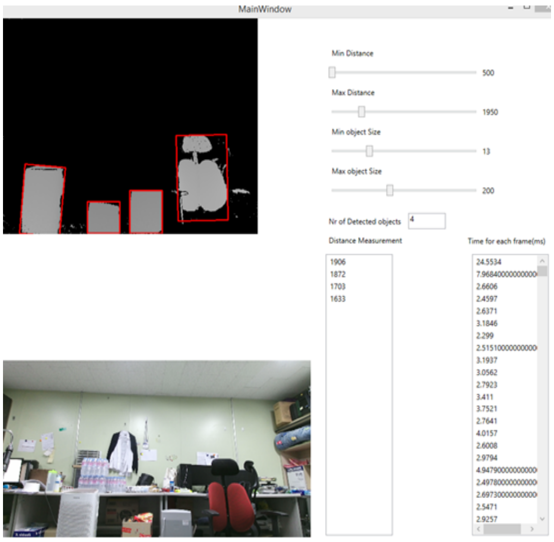

Fig. 3 Detected objects with minimum size set to 13

On Fig.3 we can identify 4 real objects within a zone delimited from 500 mm to 1950 mm instead of all kinect’s field of view as we will discuss later the importance of detecting objects in a specific area based on the distance one wants to avoid between the sensor and the detected objects. Another parameter here is the processing time which has been minimized from an average of 7 ms (as shown in Fig. 2) to an average of 2 ms (as shown in Fig.3).

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.

Table. 1 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

Table 1 shows different values recorded frame after another during 30 first frames for one detected object and the maximum variation is 8mm.

Table. 2 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

Also in Table 2, same scenario is repeated but in another interval to assess also the impact of object location and the maximum variation is 9mm.

Table. 3 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 these tables (Table 1,Table 2 and Table 3) 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.

Ⅴ. CONCLUSION

In this paper, the purpose was to use kinect depth sensor, a low cost sensor and its capabilities for essential functions in computer vision, object detection and distance measurement. We tried to detect in a very accurate way the presence of objects in front of Kinect sensor taking into 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. Results from kinect camera are promising even though there is a variation in distance from the same detected object and so many systems may use this depth sensor such as in parking area or vehicles to understand its surrounding by detecting present objects and measuring how far they are for further processing depending on application type.

ACKNOWLEDGMENTS

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] S. Nissimov, J. Goldberger and V. Alchanatis, "Obstacle detection in a greenhouse environment using the Kinect Sensor," *Journal of Computers and Electronics in Agriculture*, vol. 113, no. 3, pp. 104-115, Feb. 2015.

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

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

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

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

[6] E. Lachat, H. Macher, T. Landes and P. Grussenmeyer, "Assessment and Calibration of a RGB-D Camera (Kinect v2 Sensor) Towards a Potential Use for Close-Range 3D Modeling," *JournalRemote Sensing* , vol. 7, no. 10, pp. 13070-13097, Oct. 2015.

[7] MIT Computer Science and Artificial Intelligence Laboratory. 3D Object Tracking using the kinect [Internet]. Available: https://people.csail.mit.edu/spillai/data/papers/cvclass-project-paper.pdf.

[8] O. Wasenmuller, M. Meyer and D. Stricker, "CoRBS:

Comprehensive RGB-D Benchmark for SLAM using Kinect v2," in *Proceeding of IEEE Winter Conference on Applications of Computer Vision*, New York, pp.1-7, 2016.

[9] R. Sivalingam, A. Cherian,Joshua, N. Walczak, N. Bird, V. Morellas, B. Murphy, K. Cullen, K. Lim, G. Sapiro and N. Papanikolopoulos, "A Multi-sensor visual tracking system for behavior monitoring of At-Risk children," in *Proceedings of IEEE International Conference on Robotics and Automation*, Saint Paul (MN,USA), pp.1345-1350, 2012.

[10] V. Filipe, H. Fernades, F. Fernandes and J. Barroso.(2012, December). Blind navigation support system based on Microsoft Kinect. *Procedia Computer Science* [Online], 14(12), pp.94-101. Available : http://www.sciencedirect.com/science/article/pii/S1877050912007739?via%3Dihub.

니욘사바에릭(Niyonsaba Eric)

2015 ~ 현재 : 동의대학교 컴퓨터공학과 석사
※관심분야 : 컴퓨터 비전

장종욱(Jong-wook Jang)

1995년 2월 부산대학교 컴퓨터공학과 박사
1987년 ~ 1995년 ETRI
2000년 2월 UMKC Post-Doc.
1995년 ~ 현재 동의대학교 컴퓨터공학과 교수
※관심분야 : 유무선통신시스템, 자동차네트워크