IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 05, Issue 08 (August. 2015), ||V4|| PP 43-46

Obtainment of Background Image Using Depth Information

Soon-kak Kwon, Dong-seok Lee
Dept. of Computer Software Engineering, Dongeui University

Abstract: Extracting objects from background using the color image has some weaknesses. Using the depth image, these weaknesses can be solved. However, the depth information has some noise so the extraction of the object is incorrect. In this paper, we measure the accuracy of the methods in obtaining the background image from accumulation images. We find the best method of obtaining the background depth image.

Keywords: - *Extraction, Depth information, Depth camera*

I. INTRODUCTION

Image processing is one of the most important issues to extract an object from a background. To extract the object, the process of separating the foreground from a background is required. This process has various methods. The method of comparing a captured image from the background image which does not have any object is one of these methods.

However, using this method with only color image has the following limits: (1) having high complexity if all channel in the color image are used; (2) having many errors if only brightness is used. As the 2010 Microsoft's Kinect has been released by the end widespread, studies in various field of the image processing using the depth camera have been increased[1,2,3]. The depth information which was obtained by the depth camera is not affected by the illumination. Thus, we can extract more correctly the foreground from the background using the depth camera than using the color camera. However, the depth information measured by the depth camera is not correct because of the noise. Thus, we have to accumulate depth images and obtain the correct background image using these images to correctly extract the object.

In this paper, we accumulate the background to create the correct background image. After that, we make a background depth image with various methods. We measure the accuracy of a background depth image on each case. We can know the best method of making a background depth image from an accumulated images through our simulation. Through this method, we can obtain more correct background depth image using the depth information than the color image.

II. OBTAINMENT OF BACKGROUND IMAGE USING DEPTH INFORMATION

We can obtain the object by comparing the captured depth image with the background depth image using the depth camera. In this method, it is important to extract correctly the object from the captured image to find a correct background image. However, the background image made by only one background depth map is not correct due the depth information obtained by the depth camera which is also incorrect because of the depth camera’s method of measuring depth information [4,5]. The noise in the background image causes a problem where the object near to background cannot be extracted. Figure 1 shows the graph showing the error of measured depth information.

In Fig.1, the measured depth information has some noise. However, this noise isn’t continuous and this depth information is changed based on the correct depth information. We can estimate the correct depth information from many depth images. Thus, the process of accumulating the depth images is required to obtain the correct depth information. To estimate the correct depth information, we need to obtain the background image from the accumulated images by a statistic method. Then, we can correct the background image using this method.

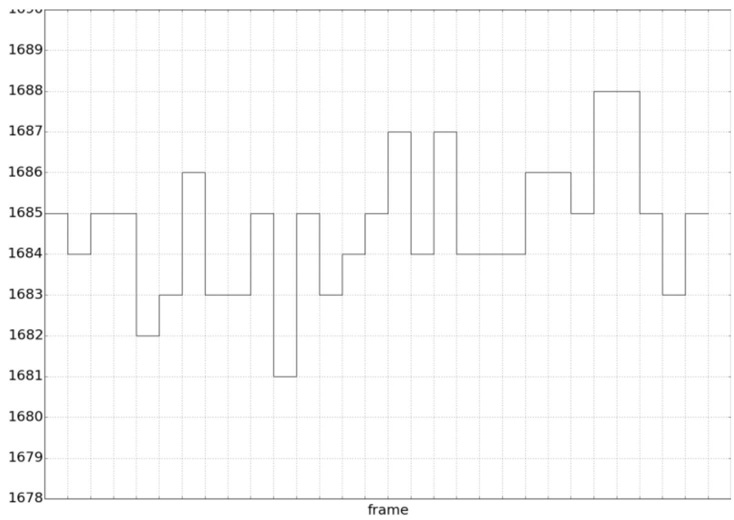

Fig. 1: The graph of changing depth information on a pixel.

We use the depth images captured a plane by the depth camera to measure accuracy of the background image. For measuring the accuracy of the background image, we capture the plane area using the depth camera. When the depth camera captures the plane area, the plane area in camera coordinate system on real world has the camera focus as the origin point and the direction of the optical axis is as same as that in front of the camera, and can be represented as a plane equation, Eq. 1.

$$ax + by + cz + d = 0 \quad (1)$$

Eq. 1, (a, b, c) means the normal vector of the plane and we can find d to substitute one point of this plane to the equation. We can obtain a normal vector $\boldsymbol{N} \equiv (a_N, b_N, c_N)$ of the plane in the camera coordinate system[6]. We also find z_0 as the depth information of the image center. We can substitute a point (0, 0, z_0) on plane to Eq. 1 because z_0 is the depth information of the image center, so we find d as - $c_N z_0$. Thus, equation 1 can be represented as Eq. 2 to substitute (a_N, b_N, c_N, - $c_N z_0$) to (a, b, c, d) in Eq. 1.

$$a_N x + b_N y + c_N (z - z_0) = 0 \quad (2)$$

The relationship between $\boldsymbol{P} \equiv (x, y, z)$ in the camera coordinate system and $\boldsymbol{P}_i \equiv (x_i, y_i)$ in the image coordinate system on the image is that the coordinate system has the image center as the origin point, and is represented as Eq. 3 and Eq. 4[7].

$$x = \frac{z}{f} x_i \quad (3)$$

$$y = \frac{z}{f} y_i \quad (4)$$

Equation 2 can represent as Eq. 5 using the relationship between the camera coordinate system and the image coordinate, Eq. 3 and Eq. 4. In this regard, z represents the depth information of the point $\boldsymbol{P}_i \equiv (x_i, y_i)$ because the depth information is the distance from XY plane composed of the depth camera. Thus Equation 5 is the relationship between the depth information and the point $\boldsymbol{P}_i \equiv (x_i, y_i)$ on the plane, or on the captured image. Figure 2 shows the above processes.

$$z = \frac{c_N z_0}{\frac{a_N}{f} x_i + \frac{b_N}{f} y_i + c_N} \quad (5)$$

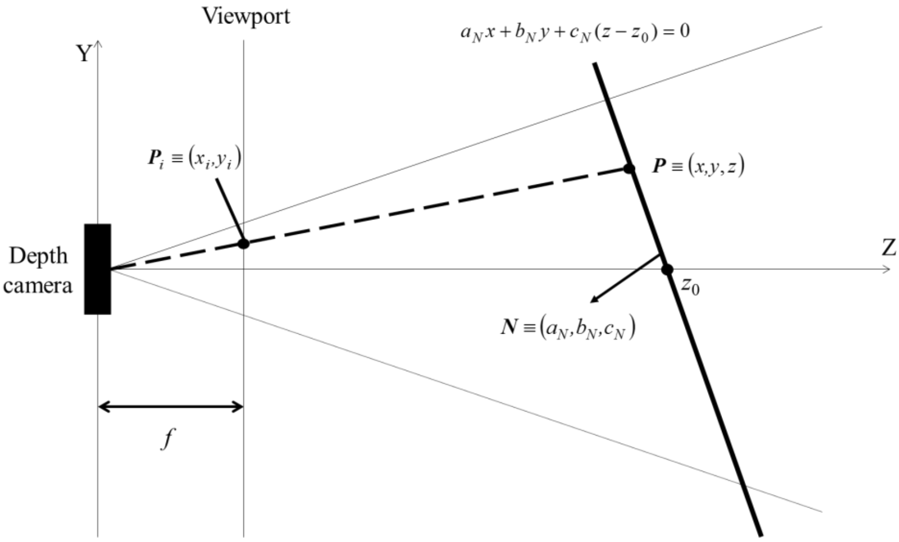

Fig. 2: The plane captured by the depth camera.

We can estimate the depth information of P_i on the plane using Eq. 5. Also, we already know the depth information to capture the plane area by the depth camera. We let the depth information that is captured by the depth camera as $\tilde{d}$ and the depth information that is estimated using Eq. 5 as $\bar{d}$. Then, we can measure the accuracy of the background image to compare between $\tilde{d}$ and $\bar{d}$. Figure 3 shows the difference between $\tilde{d}$ and $\bar{d}$. In Fig. 3, the dash line is $\bar{d}$ and the solid line is $\tilde{d}$.

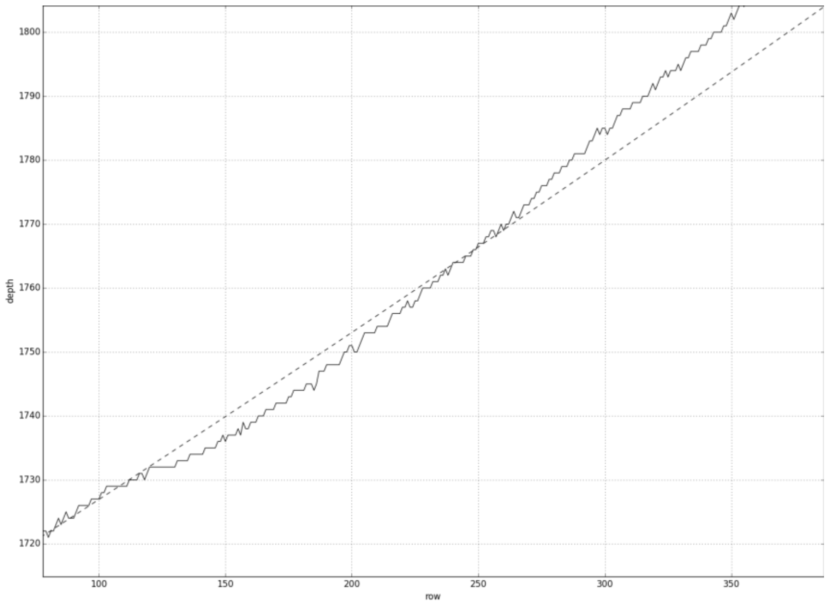

Fig. 3: The difference between the depth information captured and estimated.

We obtain the background image from the accumulated images by various statistic methods. We use the average, the mode and the median of the accumulated images to obtain background image. In this regard, we calculate variance of the difference between $\tilde{d}$ and $\bar{d}$ in each pixel in the background image using Eq. 6 to measure the accuracy of the background image.

$$\sigma^2 = \frac{1}{width \times height} \sum_{x_i, y_i} (\tilde{d} - \bar{d})^2 \quad (6)$$

We measure the accuracy of the background image using each method of obtaining the background image. We also measure by changing the number of accumulated images n. The result of measure is shown Table 1.

Table 1: The result of measure accuracy of the background image.

N	The average	The median	The mode
100	92.83	97.51	96.30
250	91.39	91.04	95.40
500	91.07	90.75	95.13
1000	90.96	90.64	94.69

We find that the accuracy is the more correct if more backgrounds are accumulated in all cases of method of obtaining the background image. We find that accuracy is the best if the median of accumulated images is used to obtain the background image. However, if the number of accumulated images is not enough, the accuracy of the method using the median drops rapidly. The accuracy of the method using the average is similar to the best in all cases of the number of accumulated images. Also, the method using the average is faster than other methods. The accuracy of the method using the mode is worse than other two methods.

In the above result, we find that it is the best to use the median as the method of obtain the background image if it can afford to accumulate enough images. If it have to use a small number of images or need fast speed, the method using the average is the best.

III. CONCLUSION

In this paper, we know that we have to accumulate backgrounds to obtain the correct background image. We find the plane equation by the normal vector obtained using the depth camera and we measure the accuracy of the background image obtained by accumulated background image to compare an estimated depth information between using the plane equation and the depth information from captured image. We know that the method using the median of depth information is the best correct. We also know that the more the correct background image we can obtain, the more background frames we accumulate.

The conventional image processing using a color image has week points like similar color, illumination, and etc. We expect that we can overcome this week points using depth image.

IV. ACKNOWLEDGMENT

This work (Grants No. C0248402) was supported by Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration in 2014.
Corresponding author: Soon-kak Kwon (skkwon@deu.ac.kr).

REFERENCES

[1] M. Siddiqui and G. Medioni, “Human pose estimation from a single view point, real-time range sensor”, In Workshop on Computer Vision for Computer Games at Conference on Computer Vision and Pattern Recognition, 2010.

[2] R. Munoz-Salinas, R. Medina-Carnicer, F.J. Madrid-Cuevas, and A. Carmona-Poyato, “Depth silhouettes for gesture recognition”, Pattern Recognition Letters, vol.29, no.3, pp.319-329, 2008.

[3] P. Suryanarayan, A. Subramanian, and D. Mandalapu, “Dynamic hand pose recognition using depth data”, In International Conference on Pattern Recognition, 2010.

[4] B. Choo, M. Landau, M. DeVore and P.A. Beling, “Statistical Analysis-Based Error Models for the Microsoft KinectTM Depth Sensor”, Sensors, vol.14, no.9, 2014.

[5] C. V. Nguyen, S.Izadi, and D. Lovell, “Modeling Kinect Sensor Noise for Improved 3D Reconstruction and Tracking”. 3D Imaging, Modeling, Processing, Visualization and Transmission, 2012 Second International Conference on, 2012.

[6] H.W. Yoo, W.H. Kim, J.W. Park, W.H. Lee, and M.J. Chung, “Local Normal Vector-based Fast Plane Detection using a Depth Camera,” The 8th Korea Robotics Society Annual Conference, pp. 15-18, 2013.

[7] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, Cambridge, UK, 2008.