Advances in 3D Camera: Time-of-Flight vs. Active Triangulation

Daesik Kim[1] and Sukhan Lee[2]

[1] Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon, Korea 440-746
daesik80@skku.edu

[2] Department of Electrical and Computer Engineering and Department of Interaction Science, Sungkyunkwan University, Suwon, Korea 440-746
lsh@ece.skku.ac.kr

Abstract. Over the last decade, numerous 3D camera techniques have been proposed and advanced dramatically. One main approach is time-of-flight (TOF) and the other is active triangulation. Each has its own strengths and weaknesses. In this paper, we overview the principle of each method and compare the advantages and disadvantages in detail, and introduce several commercially available 3D cameras and their characteristics.

1 Introduction

A demand for real-time range images has increased explosively in the computer vision and robotic community. The advances in 3D camera have created a new research area in many applications, such as gesture recognition, human/object tracking, 3D SLAM, 3D scene reconstruction, and mixed/augmented reality etc. The principle of the 3D camera can be largely classified into time-of-flight (TOF) and active triangulation. Although TOF and active triangulation had become obsolete technologies, they have come in the limelight again recently, because of their potential ability to measure the 3D scene with a high frame rate and high spatial and depth resolution. In this paper, we overview the principles of 3D camera techniques, and analyze the advantages and disadvantages of each method.

This paper is organized as follows. The principles and characteristics of the time-of-flight and active triangulation technologies are described in section 2 and 3, respectively. The specifications of the commercially available sensors and their pros and cons are summarized in section 4. Finally section 5 concludes the work.

2 Time-of-Flight

2.1 Pulse Runtime Measurement

The basic principle of the TOF camera is to measure the travel time when the pulsed light is emitted, reflected by an object, and received as shown in Fig. 1 (left). Since the speed of light is known, the distance can be computed as

$$d = \frac{t}{2} \cdot c, \tag{1}$$

S. Lee et al. (Eds.): Intelligent Autonomous Systems 12, AISC 193, pp. 301–309.
springerlink.com

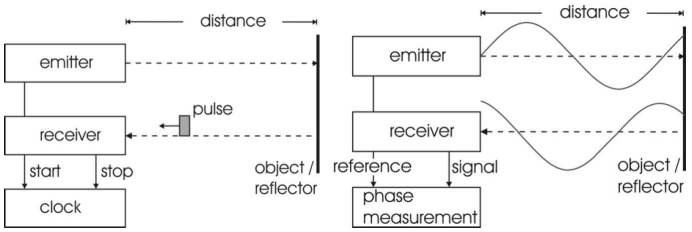

Fig. 1. Time-of-flight distance measurement principle. Pulse Runtime (left) and Phase Shift (right). [1]

where d is the distance between the sensor and the object; t is the travel time between emitting and receiving; c is the speed of light ($c \approx 3 \cdot 10^8 m/s$).

Although the concept looks simple, there are several challenges to achieve higher accuracy; high-accuracy clock time measurement and the short light pulse generation with high repetition rates should be overcome.

2.2 Phase Shift Measurement

In order to avoid high precision clocks and short light pulse, an alternative approach is to measure the phase shift between the emitted and received sinusoidal modulated light as shown in Fig. 1 (right) [1–6].

Let the emitted sinusoidal signal $g(t)$ and received signal $s(t)$ be as follows:

$$\begin{aligned} g(t) &= cos(\omega t), \\ s(t) &= b + acos(\omega t + \phi), \end{aligned} \tag{2}$$

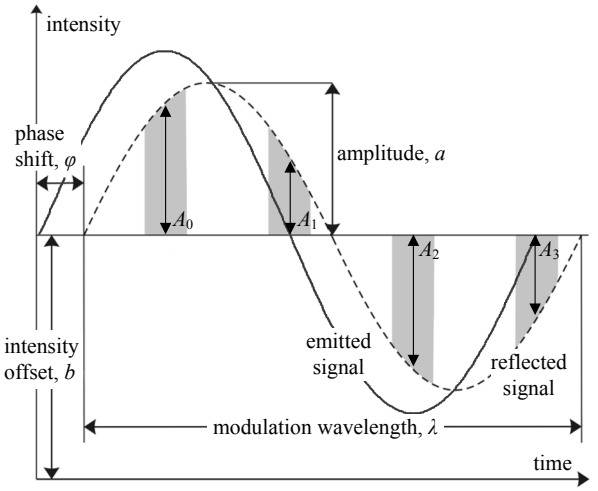

Fig. 2. Phase shift distance measurement principle

where ω is the modulation frequency; a is an amplitude; b is an intensity offset; φ is a phase shift as shown in Fig. 2. In order to retrieve the phase shift φ, it should be demodulated by cross-correlation of received signal with emitted signal:

$$c(\tau) = s * g = \int_{\infty}^{\infty} s(t) \cdot g(t+\tau) dt, \tag{3}$$

where τ is the offset. This function can be simplified as

$$c(\tau) = \frac{a}{2} cos(\omega\tau + \varphi) + b. \tag{4}$$

Let the four measurements be as follows:

$$A_i = c(i \cdot \frac{\pi}{2}), \qquad i = 0, \cdots, 3 \tag{5}$$

then the unknown a, b, and φ can be obtained as follows:

$$\begin{aligned} a &= \frac{\sqrt{(A_3 - A_1)^2 + (A_0 - A_2)^2}}{2}, \\ b &= \frac{A_0 + A_1 + A_2 + A_3}{4}, \\ \varphi &= arctan2(A_3 - A_1, A_0 - A_2). \end{aligned} \tag{6}$$

Finally, the distance between the sensor and the object can be computed as follows:

$$d = \frac{c}{4\pi\omega}\varphi, \tag{7}$$

where c is the speed of light.

The drawback of this approach is the light should be integrated over time to improve the measurement accuracy and to reduce the noise, so this limits the frame rate of the sensor. Moreover, it is hard to obtain a high-resolution depth image.

2.3 Intensity Measurement by a Shuttered Sensor

Another approach is to measure the time of flight with a fast image shutter [7–11]. The concept is to measure the accumulated light intensity instead of the pulsed light runtime. The emitter generates an IR laser pulse (*light wall*) and the receiver accumulates the reflected light carrying an imprint of the objects. The depth information can now be extracted from the reflected *deformed light wall*, by deploying a fast image shutter in front of the CCD chip, and blocking the incoming light, as shown in Fig 3.

In practice, objects have various reflectivity coefficients, so this should be compensated. By dividing the front portion of intensity $I_{shutter}$ (see Fig. 3 (d)) by the corresponding portion of the total intensity I_{total} (see 3 (c)), the ratio, which indicates the normalized depth, can be obtained as follows:

$$\alpha = I_{shutter} / I_{total}. \tag{8}$$

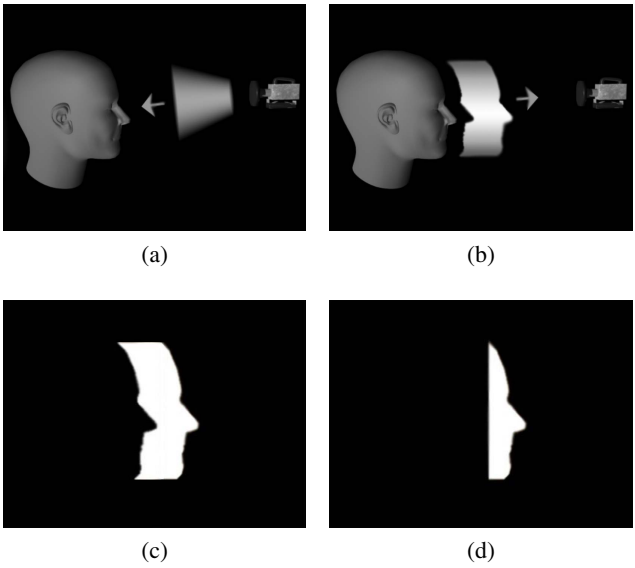

Fig. 3. The principle of the intensity measurement approach by the shutter sensor: (a) A *light wall* is emitted from the sensor. (b) It is reflected by the object. (c) The total reflected light from the object. (d) The front portion of the light is cut by a shuttered sensor. If the object distance is close, the front portion of the reflected light is big, otherwise small [7].

The measurable minimum and maximum distances are as follows:

$$\begin{aligned} d_{min} &= \frac{t_{min} \cdot c}{2}, \\ d_{max} &= \frac{t_{max} \cdot c}{2}, \end{aligned} \tag{9}$$

where c is the speed of light; t_{min} and t_{max} is the time when the shutter opens and closes, respectively. The absolute distance is

$$d = d_{max} - (d_{max} - d_{min})\alpha. \tag{10}$$

Ideally, α is equal when the same depth is measured regardless of the objects' reflectivity, but in practice, this is not true [11]. Therefore, this should be taken care of.

Theoretically, this overcomes the drawback of the pulse run-time and phase shift measurement approach, and can achieve the high-resolution depth image; however, not many research results are available.

3 Triangulation

3.1 Point or Line Light Measurement by a Camera

The basic geometry for an active triangulation system is shown in Fig. 4a. Here, it is assumed that the origin of the reference frame is attached to the center of projection of

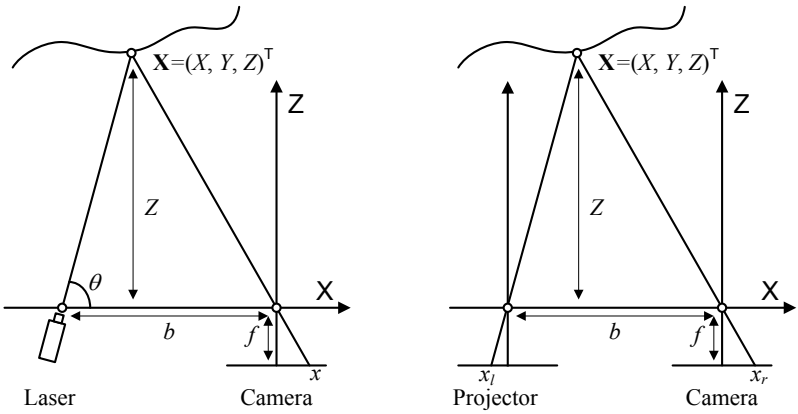

Fig. 4. The principle of active triangulation

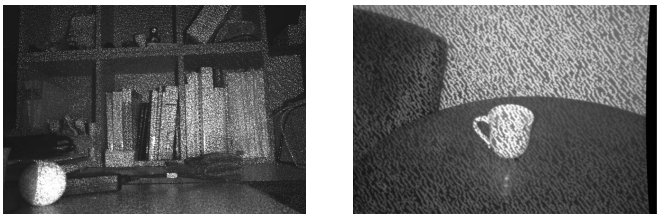

Fig. 5. Coded Pattern: (a) Kinect's pattern. (b) Willow Garage's pattern [12].

a camera. Let f be the focal length and b be the baseline, which is a distance between the line (or point) light source (e.g., laser) and the origin. In this setup, the coordinates of a point $\mathbf{X} = (X,Y,Z)^{\mathsf{T}}$ are given by

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{b}{f\cot\theta - x} \begin{bmatrix} x \\ y \\ f \end{bmatrix}, \tag{11}$$

where x and y is the corresponding image point.

Since this is a line (or point) measurement system, in order to measure the area, the angle of the light source should be scanned (rotated). Therefore, this approach is not suitable for moving objects.

3.2 Coded Pattern Measurement by a Camera

Fig. 4b and Fig. 5a show the principle of another simple active triangulation, which is composed of an imaging device (e.g., camera) and a projecting device (e.g., projector) [13, 14].

The difference between the projector and the camera is the direction of the projected light. A projector is a device to project the light to outside, but a camera is a device to acquire the light from outside. Since the projector model is the same as the camera model, the principle of the active triangulation becomes the same as the conventional passive stereo system. Here, it is assumed that there is no rotation between two devices. Let the focal length, f, the baseline, b, which is the distance between two devices, and the disparity $d = x_l - x_r$, then the distance Z can be estimated by the triangle similarity:

$$\frac{b}{Z} = \frac{b-d}{z-f}, \tag{12}$$

$$Z = \frac{fb}{d}. \tag{13}$$

X and Y are also computed as follows:

$$X = \frac{xZ}{f} \quad \text{and} \quad Y = \frac{yZ}{f}. \tag{14}$$

The distance Z is proportional to the focal length, f, and the baseline, b, but inversely proportional to the disparity.

3.3 Coded Pattern Measurement by Stereo Camera

The system composed of two cameras and one projector is another approach [12]. In order to robustly solve the correspondence problem between two cameras, the pattern is coded so that every block is well discriminated from every other block. The projector is for generating the texture to the objects and the conventional block matching algorithm is used for a stereo camera as shown in Fig. 5b.

3.4 Depth Resolution

The depth resolution of the triangulation can be computed by differentiating the equation of the triangulation (13):

$$\frac{\partial Z}{\partial d} = -f\frac{b}{d^2}. \tag{15}$$

The disparity is derived from the equation (13) as follows:

$$d = \frac{fb}{Z}. \tag{16}$$

By plugging equation (16) into (15),

$$\frac{\partial Z}{\partial d} = -\frac{Z^2}{fb}. \tag{17}$$

Since the equation (17) means the rate of change of Z with respect to d, if we denote $\partial Z = \Delta Z$ and $\partial d = \Delta d$, and take the absolute value to make the resolution positive, the depth resolution is

$$\Delta Z = \frac{Z^2}{fb}\Delta d. \tag{18}$$

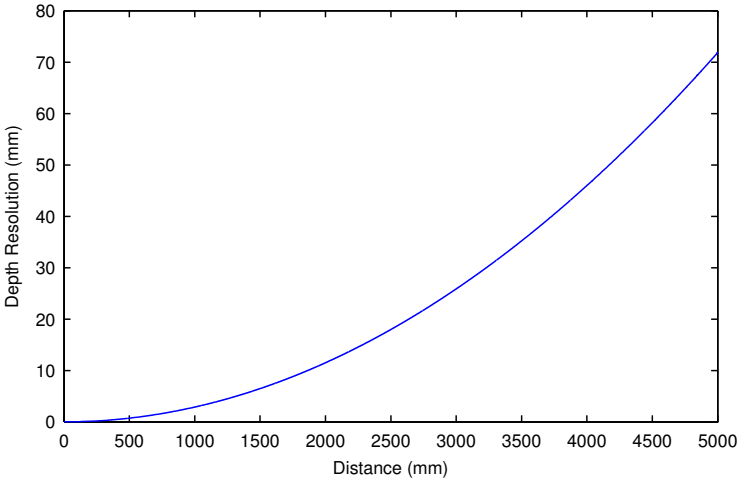

Fig. 6. The depth resolution of the Kinect

From equation (18), we can infer that the depth resolution becomes higher (i.e., ΔZ becomes smaller.) if the focal length f and baseline b become bigger and the distance Z becomes shorter. Fig. 6 shows the depth resolution of the Kinect.

Table 1. TOF-Based 3D Camera Specification

	SR4000 (Mesa)	CamCube (PMD)
Light Source	LED (NIR)	LED (NIR)
Image Resolution	176×144	200×200
Measurement Range	0.8–5.0m	0.3–7.0m
Field of View	$43^\circ \times 34^\circ$ (1) $69^\circ \times 56^\circ$ (2)	$40^\circ \times 40^\circ$
Frame Rate	50fps	40 (5), 60 (6), 80 (7) fps
Size	65×65×68mm (3) 65×65×76mm (4)	194×60×60mm (8)
Repeatability (1σ)	4mm @ 2m (1) 6mm @ 2m (2)	3mm @ 4m

(1) Standard Field of View, (2) Wide Field of View
(3) USB Camera, (3) Ethernet Camera
(5) 200×200, (6) 176×144, (7) 160×120 pixels
(8) One camera and two illumination units

4 Specification Comparison

In this section, we compared the tow commercially available TOF-based 3D cameras and two active triangulation-based 3D camera (see Table 1 and 2).

Table 2. Active Triangulation-Based 3D Camera Spec

	Kinect (X-tion)	Artec 3D Scanner
Light Source	LED (NIR)	Flash Bulb
Image Resolution or X-Y Resolution*	640×480	0.5mm*(5) 1.0mm*(6)
Measurement Range	0.8–3.5m	0.4–1.0m*(5) 0.8–1.6m*(6)
Field of View	58×45°	30×21°(5) 41×32°(6)
Frame Rate	30fps(1) 60fps(2)	15fps
Size	283×39×62mm(3) 180×35×50mm(4)	72×222×120mm(5) 70×353×114mm(6)
Repeatability(1σ) or Accuracy**	1.5mm @ 1m 6mm @ 2m	0.1mm @ 1m**(5) 0.2mm @ 1m**(6)

(1)640×480, (2)320×240
(3)Kinect, (4)X-tion, (5)Artec M, (6)Artec L
*3D X-Y Resolution (up to)
**3D Point Accuracy (up to)

In summary, the advantages of the TOF-based 3D Camera are that it is compact and accurate at a relatively long distance. The disadvantage is that the image resolution is very low. On the other hand, the active triangulation approach is more accurate in the near range measurement and it is able to provide a high-resolution image. The weakness is that in order to measure the relatively long distance, the baseline should be longer to preserve the accuracy. This makes the sensor size big.

5 Conclusion

In this paper, we overview the principles of the 3D camera techniques, and analyze the advantages and disadvantage of each method.

Acknowledgment. This research was performed for the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs (F0005000-2010-32), and in part for the KORUS-Tech Program (KT-2008-SW-AP-FSO-0004) funded by the MKE, Korea. This work was also partially supported by the MEST, Korea, under the WCU Program supervised by the KOSEF (R31-2008-000-10062-0), and by MKE, Korea under ITRC NIPA-2010-(C1090-1021-0008)(NTIS-2010-(1415109527)).

References

1. Kahlmann, T., Remondino, F., Ingensand, H.: Proceedings of the ISPRS Commission V Symposium 'Image Engineering and Vision Metrology', pp. 136–141 (2006)
2. Lange, R.: 3D time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology. Ph.D. thesis, University of Siegen (2000)
3. Oggier, T., Lehmann, M., Kaufmann, R., Schweizer, M., Richter, M., Metzler, P., Lang, G., Lustenberger, F., Blanc, N.: Proceedings of SPIE: Specific Applications: Sensors and Medical Optics
4. Blanc, N., Oggier, T., Gruener, G., Weingarten, J., Codourey, A., Seitz, P.: Proceedings of IEEE Sensors, pp. 471–474 (2004)
5. Oggier, T., Lustenberger, F., Blanc, N.: Miniature 3D TOF Camera for Real-Time Imaging. In: André, E., Dybkjær, L., Minker, W., Neumann, H., Weber, M. (eds.) PIT 2006. LNCS (LNAI), vol. 4021, pp. 212–216. Springer, Heidelberg (2006)
6. Gokturk, S., Yalcin, H., Bamji, C.: Conference on Computer Vision and Pattern Recognition Workshop, p. 35 (2004)
7. Iddan, G., Yahav, G.: Proceedings of SPIE: Videometrics and Optical Methods for 3D Shape Measurements, pp. 48–55 (2001)
8. Gvili, R., Kaplan, A., Ofek, E., Yahav, G.: Proceedings of SPIE Video-Based Image Techniques and Emerging Work, pp. 564–574 (2003)
9. Medina, A., Gaya, F., del Pozo, F.: Journal of the Optical Society of America A 23(4), 800 (2006)
10. Yahav, G., Iddan, G., Mandelboum, D.: Digest of Technical Papers of International Conference on Consumer Electronics, pp. 1–2 (2007)
11. Davis, J., Gonzalez-banos, H.: IEEE International Workshop on Projector-Camera Systems (2003)
12. Konolige, K.: Proceedings of IEEE International Conference on Robotics and Automation, pp. 148–155 (2010)
13. Sazbon, D., Zalevsky, Z., Rivlin, E.: Pattern Recognition Letters. 26(11), 1772 (2005)
14. García, J., Zalevsky, Z., García-Martínez, P., Ferreira, C., Teicher, M., Beiderman, Y.: Applied Optics 47(16), 3032 (2008)