Error correction of depth images for multiview time-of-flight vision sensors

Yu He¹ and Shengyong Chen²

Abstract
The developing time-of-flight (TOF) camera is an attractive device for the robot vision system to capture real-time three-dimensional (3D) images, but the sensor suffers from the limit of low resolution and precision of images. This article proposes an approach to automatic generation of an imaging model in the 3D space for error correction. Through observation data, an initial coarse model of the depth image can be obtained for each TOF camera. Then, its accuracy is improved by an optimization method. Experiments are carried out using three TOF cameras. Results show that the accuracy is dramatically improved by the spatial correction model.

Keywords
Robot vision, 3D image, TOF, multiview, depth camera, error correction

Introduction
The three-dimensional (3D) vision is vital for a robot system working in uncertain environments. In recent years, several novel 3D sensing devices have been developed, and the computational methods are proposed accordingly. Among these devices, the time-of-flight (TOF) camera receives wide attention. However, its resolution and sensing range are limited. Compared with other types of depth sensors, the advantage of TOF camera is real-time depth acquisition of the entire 3D scene.¹,²

The raw depth data directly obtained by the TOF camera usually suffer from systematic and nonsystematic errors, as is shown in Figure 1. Systematic errors include integral time error,³,⁴ amplitude ambiguity error,⁵ and temperature drift error.⁶ Nonsystematic errors include light scattering error,⁷ boundary ambiguity error,⁸ multipath reflection error,⁹ phase wiggling error,¹⁰ and motion blur error.¹¹ The systematic errors are generally determined by the design principle and component accuracy of the TOF camera.¹² Such errors occur frequently and have relatively single error characteristics. Theoretically, systematic errors should be eliminated through calibration after the cameras are manufactured. However, in the application of visual system,¹³,¹⁴ due to the limitation of calibration technology and the influence of uncertainties in calibration process, the raw data acquired by TOF cameras are still subject to a small amount of systematic errors. In contrast, the nonsystematic errors occur randomly, and different scenes, target objects, and measurement distances cause different error influences. Therefore, the error correction algorithm needs to consider specific measurement conditions.

According to the acquisition methods of standard data, the systematic error correction methods can be divided into two parts, which are direct calibration methods and indirect...
Previous approaches of TOF camera calibration adopt linear methods to measure the distance errors using high precision measurement equipment. Instead, this article presents an attempt to obtain a spatial distance error model of a TOF camera, which benefits the depth image optimization. The principle of the TOF camera is briefly described in the second section. Then, we address how to determine the error model of the TOF camera and how to optimize the depth image based on multicamera system in the third section. Experiments are performed on two different scenes so as to evaluate the performance of the optimization method in the fourth section. Our conclusions are summarized in the fifth section.

**Depth computation principle**

The depth computation principle introduced in this article is the phase-shift measurement. An infrared (IR) light is emitted to the target object via light-emitting diodes, and the TOF sensor detects the reflected IR component. The depth can be calculated by measuring the phase shift between the radiated and reflected IR signals. Four control signals with electric charge values are used to calculate the phase shift. As shown in Figure 2, there is a 90° phase delay between each control signal. \( \phi \) is defined as

\[
\phi = \arctan\left( \frac{C_3 - C_4}{C_1 - C_2} \right)
\]

(1)

where \( C_1, C_2, C_3, \) and \( C_4 \) represent the amount of electric charge of the control signals. Then, the distance \( D \) can be calculated by

\[
D = \frac{c}{2f} \frac{\phi}{2\pi}
\]

(2)

where \( c \) is the speed of light and \( f \) is the frequency of the signal.

**Modeling and correction**

*Representation of error model*

Distance measurements of a single TOF camera always accompany systematic error, as is shown in Figure 1. In this article, we focus on mixed errors of systematic and nonsystematic, the low accuracy of TOF data, especially the depth information corresponding to the pixels near the border of the projection plane. To obtain the spatial error distribution of the camera measurement and ensure the accuracy of the error model establishment, the experimental object for error modeling is a flat and untextured wall surface (as is shown in Figure 3(a)). Figure 4 shows the point cloud of the plane obtained by the TOF camera. According to the vertical distance between the camera and plane, the error value of each point in the plane point cloud can be calculated. Figure 5 shows the error distribution of plane measurement.
The raw depth images taken by the TOF camera are inaccurate, especially at the edges of the images. According to the error distribution and the periodic characteristics of error variation, the spatial error model of TOF camera is constructed as follows:

\[
\delta(d) = c_0 + d \cdot c_1 \sin(c_2 \cdot d + c_3)
\]  

where \(\delta(d)\) is the pixel measured distance, \(c_0\) is a constant, \(c_1\) is the scale factor, \(\sin(c_2 \cdot d + c_3)\) is a "wiggling error," \(c_2\) is the angular frequency, and \(c_3\) is the phase shift of sinusoidal function.

\[
d(x, y) = \sqrt{m^2 + x^2 + y^2}
\]  

where \(d\) is the pixel measured distance, \(m\) is the depth value of the center pixel, \(c_0\) is a constant, \(c_1\) is the scale factor, \(\sin(c_2 \cdot d + c_3)\) is a "wiggling error," \(c_2\) is the angular frequency, and \(c_3\) is the phase shift of sinusoidal function.

Figure 2. Depth is calculated by the phase difference between the radiated and reflected IR signals, and the quantities \(C_1\) to \(C_4\) are the amount of electric charge of the control signals \(S_1\) to \(S_4\). IR: infrared.

Figure 3. Depth data acquisition. (a) Flat surface acquisition and (b) sphere surface acquisition.
To estimate the parameters, we find a fitting surface and minimize the distance between the error value to the surface:

$$E = \frac{1}{N} \sum_{i=1}^{N} (\delta(d(x_i, y_i)) - z_i)^2$$  \hspace{1cm} (5)

where $E$ is the loss function of the distance, $N$ is the number of pixels, and the parameter estimation can be done by minimizing the loss function $E$.

**Depth correction**

There are some contributions available for depth map denoising and resolution improvement, but they often perform not well for this newly developed device. In this article, we report an optimization method to improve the accuracy of the depth map. Figure 6 shows the most basic multi-TOF camera system. The depth image with error captured by the mid camera can be corrected by the multicamera system. The specific implementation process is given in Table 1.

As is shown in Figure 6, $A$, $B$, and $C$ are three different points on the object surface. $A_1$, $A_2$, and $A_3$ are the projection pixels of $A$ in the mid, left, and right camera views, respectively. $[x, y, z]$ is the coordinate of $A$ in the World Coordinate System based on the projection rules of TOF camera, as in formula (6)

$$\begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = L_M \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$  \hspace{1cm} (6)

where $[u_1, v_1]$ is the coordinate of $A_1$ on the projection plane of the mid camera. $L_M$ is the projection matrix of the mid camera.

In the error correction experiment, the proposed method faces two situations: (1) the correction point can be captured by the TOF cameras in three views and (2) the correction point can be captured by the mid camera and the camera in one side view, however, the camera on the other side view cannot capture the point.

In the first case, we take point $A$ as an example, which can be captured by cameras in three views. Firstly, we calibrate three cameras in pairs. Based on the transform matrix calculated from calibration, the corresponding pixels of point $A$ on left and right camera projection plane can be determined, which are $A_2[u_2, v_2]$ and $A_3[u_3, v_3]$. To check occlusion, we define hypothetical depths for each point in different camera views. $H_{A_2}$ and $H_{A_3}$ are the hypothetical depths of point $A$ in left and right camera views. The hypothetical depths can be calculated from the measurement depth in the mid camera view based on transform matrices. In this case

$$\begin{cases} H_{A_2} = D_{A_2} \\ H_{A_3} = D_{A_3} \end{cases}$$  \hspace{1cm} (7)

where $D_{A_2}$ and $D_{A_3}$ are the measurement depths in left and right camera views.

Secondly, we calculate the correction weights of the left camera and right camera according to their error models

$$w_L = \frac{\delta(u_2, v_2)}{\sum_{i=1}^{m} \sum_{j=1}^{n} \delta(u_i, v_j)}$$  \hspace{1cm} (8)
\[
W_M = \frac{\delta(u_1, v_1)}{\sum_{i=1}^{m} \sum_{j=1}^{n} \delta(u_i, v_j)} \quad (9)
\]
\[
W_R = \frac{\delta(u_3, v_3)}{\sum_{i=1}^{m} \sum_{j=1}^{n} \delta(u_i, v_j)} \quad (10)
\]

where \([m, n]\) is the pixel size of projection plane. Finally, the corrected depth of point A can be calculated as
\[
W_L = 1 - \frac{w_L}{w_L + w_M + w_R} \quad (11)
\]
\[
W_M = 1 - \frac{w_M}{w_L + w_M + w_R} \quad (12)
\]
\[
W_R = 1 - \frac{w_R}{w_L + w_M + w_R} \quad (13)
\]
\[
D_{cA} = W_L \cdot D_{A2} + W_M \cdot D_{A3} + W_R \cdot D_{A4} \quad (14)
\]

where \(D_{A1}, D_{A2},\) and \(D_{A4}\) are the corresponding depth values of point A in different TOF cameras, and \(D_{cA}\) is the corrected depth value.

In the second case, we take point B as an example. In this case, point B is shielded by point A in the view of the right camera. We compare the hypothetical depths to the measured depths in the corresponding camera view. If the hypothetical depth is bigger than the measured depth, we set the depth of this point in the corresponding camera view equals to zero, otherwise, we reserve the measurement depth. For point B
\[
\begin{cases}
H_{B_L} = D_{B_L} \\
H_{B_R} > D_{B_R}
\end{cases} \quad (15)
\]
\[
D_{cB} = W_L \cdot D_{B2} + W_M \cdot D_{B3} + W_R \cdot 0 \quad (16)
\]

where \(H_{B_L}\) and \(H_{B_R}\) are the hypothetical depths of point B in the left and right camera views, \(D_{B_L}\) and \(D_{B_R}\) are the measured depths, and \(D_{cB}\) is the corrected depth value of point B.

**Experiments and results**

**Experiment settings**

Our experiments are carried out on two types of scenes. The first one is a flat and untextured wall surface (Figure 3(a)), which focuses on determining the spatial error model and verifying the performance of the correction algorithm on a flat surface. The second scene is a spherical surface, which focuses on verifying the performance of the correction algorithm on free-form surface (Figure 3(b)). The measurement results of the TOF camera are influenced by internal temperature. To achieve distance measurement stability, the SR-4000 cameras are warmed up for 40 min according to the test in the previous work.

In the first experiment scene, TOF camera is set at 1.5 m from the flat surface based on the standard working range of SR-4000 (0.3–5 m). To ensure the accuracy of the error model, the position of the TOF camera is precisely determined by a longitudinal hammer and a laser rangefinder, and the camera optical axis is perpendicular to the flat surface. In the experiment, the acquisition frequency of the TOF camera is 30 Hz. Five frames of data are collected for each measurement, and the average value is taken to ensure the stability of the data. According to the error model calculation method proposed in the section “Representation of error model,” error model parameters obtained by the least square method are presented in Table 1. Figure 7 shows the fitting surface of the error model which is in blue, and the purple points are the depth error distribution of the flat surface.

In the second experiment scene, the radius of the sphere is 0.1 m, and TOF camera is set at 1.5 m from the center of the sphere, therefore, the depth correction method can be performed with the error model calculated from the first scene directly. To avoid mutual interference between multiple TOF cameras, the depth acquisition frequency of three cameras is set as 29, 30, and 31 Hz, respectively. In the experiment, the centers of three cameras are on a straight line, spaced 0.5 m apart. Moreover, the point cloud directly acquired by the TOF camera contains jump edge points, and we use line-of-sight method to eliminate them before the depth error correction.

**Results**

According to the experiment set in the previous section, the first experiment scene obtains the following results. Figure 8 shows the comparison of depth errors of the flat surface before and after correction, where the abscissa is the number of pixels in the TOF image. The resolution of SR-4000 is \(176 \times 144\), so each point cloud contains 25,344 points. The broken line in Figure 8 is the ascending order of point cloud depth errors. Based on the precision of SR-4000 (0.01 m), we define that points with the error value greater than 0.01 m are noises. In raw TOF data, the number of noises is 7906, accounting for 31.2%. Tables 2 and 3 present the decrease of average error after depth correction. The performance of error reduction is obvious, especially for the flat surface. Moreover, in terms of noise rate reduction, our method also performs well. Figure 9
shows the depth error distribution before and after correction in $X$–$Y$ view, and errors reduce almost 50\% at corners of the image edge, which are labeled in yellow. Sphere surface depth error correction in $X$–$Z$ view is shown in Figure 10. The red line is the outline of raw TOF data, and the blue line is the outline after correction, which is closer to ground truth apparently. A more detailed comparison of 3D error distribution is shown in Figure 11, compared to the red points in Figure 11(a), the blue points in Figure 11(b) are closer to the sphere surface.

We compare our spatial correction method to the calibration method in the literature\(^3\) (green line in Figure 12) and the distance overestimation error correction method in the literature\(^4\) (purple line in Figure 12). We perform the experiments using the flat surface and sphere surface raw point cloud. As is shown in Figure 12, our method is more efficient than the others. The average error and noise rate contrasts are displayed in Tables 2 and 3.

**Table 2.** Depth correction experiment on flat surface by different methods.

|         | Average error (m) | Rate  |
|---------|-------------------|-------|
| Raw     | 0.0142            |       |
| DOEIC   | 0.0105            | 26.1\%|
| Calibration | 0.0078            | 45.1\%|
| Ours    | 0.0046            | 67.6\%|

|         | Noise number     | Rate  |
|---------|------------------|-------|
| Raw     | 6553             |       |
| DOEIC   | 4072             | 12.82\%|
| Calibration | 2667             | 20.08\%|
| Ours    | 1037             | 28.51\%|

DOEC: distance overestimation error correction.

**Table 3.** Depth correction experiment on sphere surface by different methods.

|         | Average error (m) | Rate  |
|---------|-------------------|-------|
| Raw     | 0.0092            |       |
| DOEIC   | 0.0085            | 7.6\% |
| Calibration | 0.0043            | 53.2\%|
| Ours    | 0.0024            | 73.9\%|

|         | Noise number     | Rate  |
|---------|------------------|-------|
| Raw     | 7906             |       |
| DOEIC   | 7035             | 3.43\%|
| Calibration | 3544             | 17.21\%|
| Ours    | 908              | 27.62\%|

DOEC: distance overestimation error correction.

**Discussion**

In this article, we propose a spatial error model of the TOF camera and implement error correction based on this model. In the process of first scene experiment, we construct the error model at 1.5 m. $D_{Model}$ represents the position, which constructs the error model. To estimate the application scope of the model ($D_{Model} = 1.5$ m), we apply this model constructed to correct TOF data obtained at different measurement distances. $D_{Correction}$ represents the position, which obtains the TOF data. As is shown in Figure 13, we apply the model ($D_{Model} = 1.5$ m) to correct TOF data obtained at three different measurement distances ($D_{Correction} = 0.5$ m, 1.5 m, 2.5 m) and the performance at position $D_{Correction} = 1.5$ m is better than at measurement distances. According to a more detailed experiment shown in Figure 14, when $D_{Correction}$ equals to $D_{Model}$, our method
has optimal correction performance. With the change of $D_{\text{Correction}}$, the correction algorithm will become unreliable, and the error model needs to be reconstructed in a new position. As shown in Figure 14, we treat 0.005 m as the upper limit of the average error after correction to evaluate the applicability of the model. The error model constructed at the position $D_{\text{Model}} = 1.5$ m can be applied to the measurement distance $D_{\text{Correction}} \in [1 \text{ m}, 2 \text{ m}]$. The scope of model application varying with error model construction position is shown in Figure 15.

### Conclusion

It is very important to correct these incorrect or inaccurate points before 3D registration and fusion for depth data from TOF cameras. In this article, a novel method to correct error points is presented. The experimental results show that the proposed method exhibits a satisfactory performance using the spatial model for error correction. The corrected point cloud of the target object is more approaching the true value although it is not perfectly matched. In
the future, we will perform some tests to further optimize the error model, thus improving the correction weights to obtain a higher compatibility with the ground truth.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Key R&D Program of China [2018YFB1305200] and the National Natural Science Foundation of China [U1509207].

ORCID iD
Yu He https://orcid.org/0000-0002-0357-681X

References
1. Chen S, Liu H, and Kubota N. 3-D vision for robot perception. *Int J Adv Robot Syst* 2019; 16(4). DOI: 10.1177/1729881419855495.
2. Dolson J, Baek J, Plagemann C, et al. Upsampling range data in dynamic environments. In: 2010 IEEE computer society conference on computer vision and pattern recognition (ed T Darrell), San Francisco, CA, USA, 13–18 June 2010, pp. 1141–1148. Los Alamitos, USA: IEEE.
3. He Y and Chen S. Recent advances in 3D data acquisition and processing by time-of-flight camera. *IEEE Access* 2019; 7: 12495–12510.
4. Lee C, Song H, Choi B, et al. 3D scene capturing using stereoscopic cameras and a time-of-flight camera. *IEEE Trans Consum Electr* 2011; 57(3): 1370–1376.
5. González-Ortega D, Daz-Pernas F, Martínez-Zarzuela M, et al. A kinect-based system for cognitive rehabilitation exercises monitoring. *Comput Methods Prog Biomed* 2014; 113(2): 620–631.
6. Kahlmann T and Ingensand H. Calibration and development for increased accuracy of 3D range imaging cameras. *J Appl Geod* 2008; 2(1): 1–11.
7. Lee S. Depth camera image processing and applications. In: 2012 19th IEEE international conference on image processing (ed E Saber), Orlando, FL, 30 September–3 October 2012, pp. 545–548. Los Alamitos, USA: IEEE.
8. Ghorpade VK, Checchin P, and Trassoudaine L. Line-of-sight-based TOF camera’s range image filtering for precise 3D scene reconstruction. In: *European conference on mobile
robots (ed T Duckett), Lincoln, UK, 2–4 September 2015.
Los Alamitos, USA: IEEE.
9. Fuchs S. Multipath interference compensation in time-of-flight camera images. In: 2010 20th international conference on pattern recognition (ed A Ercelebi), Istanbul, Turkey, 23–26 August 2010, pp. 3583–3586. Los Alamitos, USA: IEEE.
10. Chiabrando F, Chiabrando R, Piatti D, et al. Sensors for 3D imaging: metric evaluation and calibration of a CCD/CMOS time-of-flight camera. Sensors 2009; 9(12): 10800–10096.
11. Fürsattel P, Placht S, Balda M, et al. A comparative error analysis of current time-of-flight sensors. *IEEE Trans Comput Imaging* 2016; 2(1): 27–41.
12. Zhang J, Chen S, Liu S, et al. Normalized weighted shape context and its application in feature-based matching. Opt Eng 2008; 47(9): 097201.
13. Wang Z, Liu S, Zhang J, et al. A spatio-temporal CRF for human interaction understanding. *IEEE Trans Circ Syst Video Technol* 2017; 27(8): 1667–1662.
14. Cai H, Liu B, Zhang J, et al. Visual focus of attention estimation using eye center localization. *IEEE Syst J* 2017; 11(3): 1320–1325.
15. Lichti DD and Rouzaud D. Surface-dependent 3D range camera self-calibration. *Proc SPIE* 2009; 7239: 72390A–72390A–10.
16. Lichti DD, Xiaojuan Qi, Ahmed T, et al. Range camera self-calibration with scattering compensation. *ISPRS J Photogram Remote Sens* 2012; 74(11): 101–109.
17. Boehm J and Pattinson T. Accuracy of exterior orientation for a range camera. *Int Arch Photogramm Remote Sens Spat Inf Sci* 2010; 38(5): 103–108.
18. Kahlmann T, Remondino F, and Ingensand H. Calibration for increased accuracy of the range imaging camera swissranger. In: *Proceedings of the ISPRS Commission V Symposium ‘Image Engineering and Vision Metrology’ 2006* (Vol. 36, pp. 136–141). Isprs.
19. Kuhnt KD and Stommel M. Fusion of stereo-camera and PMD-camera data for real-time suited precise 3D environment reconstruction. In: 2006 IEEE/RSJ international conference on intelligent robots and systems (ed Y Liu), Beijing, China, 9–15 October 2006, pp. 4780–4785. Los Alamitos, USA: IEEE.
20. Lindner M and Kolb A. Lateral and depth calibration of PMD-distance sensors. Berlin: Springer, 2006.
21. Ringbeck T and Hagebeuker B. A 3D time of flight camera for object detection. Siegen: PMD Technologies GmbH, 2007.
22. Cui Y, Schuon S, Thrun S, et al. Algorithms for 3D shape scanning with a depth camera. *IEEE Trans Pattern Anal Mach Intell* 2012; 35(5): 1039–1050.
23. Kolb A, Barth E, Koch R, et al. Time-of-flight cameras in computer graphics. *Comput Graph Forum* 2010; 29(1): 141–159.
24. Fürsattel P, Placht S, Balda M, et al. A comparative error analysis of current time-of-flight sensors. *IEEE Trans Comput Imaging* 2015; 2(1): 27–41.
25. Hussmann S, Knoll F, and Edeler T. Modulation method including noise model for minimizing the wiggling error of TOF cameras. *IEEE Trans Instrum Meas* 2013; 63(5): 1127–1136.
26. Pratt V. Direct least-squares fitting of algebraic surfaces. *Proc Siggraph* 1987; 21(4): 145–152.
27. Mure-Dubois J and Hugli H. Optimized scattering compensation for time-of-flight camera. In: *Two- and three-dimensional methods for inspection and metrology V* (ed PS Huang), Vol. 6762, p. 67620. Bellingham, USA: SPIE-Int Society for Optics Engineering.
28. Yasutomi K, Usui T, Han SM, et al. A submillimeter range resolution time-of-flight range imager with column-wise skew calibration. *IEEE Trans Elect Dev* 2015; 63(1): 182–188.
29. Albota MA, Heinrichs RM, Kocher DG, et al. Three-dimensional imaging laser radar with a photon-counting avalanche photodiode array and microchip laser. *Appl Opt* 2002; 41(36): 7671–7678.
30. Li H, Liu H, Cao N, et al. Real-time RGB-D image stitching using multiple Kinects for improved field of view. *Int J Adv Robot Syst* 2017; 14(2). DOI: 10.1177/1729881417695560.
31. Mure-Dubois J and Hugli H. Fusion of time of flight camera point clouds. In: *Workshop on multi-camera and multi-modal sensor fusion algorithms and applications* (eds A Cavallaro and H Aghajan), Marseille, France, October 2008, pp. 1–12.
32. Naik N, Kadambi A, Rhemann C, et al. A light transport model for mitigating multipath interference in time-of-flight sensors. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Boston, MA, USA, 7–12 June 2015, pp. 73–81. Los Alamitos, USA: IEEE.
33. Hua K, Lo K, and Frank Wang YF. Extended guided filtering for depth map upsampling. *IEEE MultiMedia* 2016; 23(2): 72–78.
34. Wang L, Fei M, Wang H, et al. Distance overestimation error correction method (DOEC) of time of flight camera based on pinhole model. In: *Intelligent computing and internet of things* (eds K Li, M Fei, D Du, Z Yang and D Yang), Berlin: Springer, 2018, pp. 281–290.