International Journal of Control and Automation
Vol. 9, No. 6 (2016), pp.277-288
http//dx.doi.org/10.14257/ijca.2016.9.6.26

Calibration and Gray-level Image Generation for the SR4500 ToF Camera

Hui Zhang[1], Xuewen Rong[1]*, Yibin Li[1], Bin Li[1], Qin Zhang[2] and Junwen Zhang[1]

[1]*School of Control Science and Engineering, Shandong University, Jinan, China*
[2]*School of Electrical Engineering, Jinan University, Jinan, China*
rongxw@sdu.edu.cn, cse_zhangq@ujn.edu.cn

Abstract

This paper proposes a novel distance calibration and gray-level image generation method for the SR4500 Time of Flight (ToF) camera. For the distance measurement, the Gaussian Process Regression (GPR) model is proposed to calibrate the range error. With this model, fewer parameters need to be estimated and the model complexity is reduced. In order to eliminate the error data caused by the object edge, the distance threshold which determined by the range data of the current pixel is selected to remove the mixed pixel. To acquire gray-level images from the amplitude data, the amplitude histogram is used to select the amplitude threshold of the reflected light, and grayscale images can be generated through normalizing the amplitude value with this threshold. The experiment results illustrate the feasibility and effectiveness of the proposed algorithms.

Keywords*: ToF camera, Gaussian Process Regression, mixed pixel, gray-level image*

1. Introduction

Monocular vision, stereo vision and panoramic vision are widely used in the field of mobile robot. Monocular vision can provide the color information, but hard to acquire the range data of the scene. Stereo vision obtains the depth information of the environment through matching the corresponding points from two images, but it has the drawback of the low frame rate [1]. Though the panoramic vision has the advantage of providing the 360-degree field of view and high efficiency, it also needs to solve the implementation of the hardware. In recent years, a new generation of active cameras, which based on the Time-of-Flight principle, has been developed. The main advantages with respect to other 3D measurement techniques are the possibility to acquire data at video frame rates and obtain 3D point clouds without scanning and from just one point of view [2]. It is a new suitable choice for the environment perception and modeling in the field of robotics.

In the 1990s, the first all-solid-state 3D-cameras based on the Time-of-Flight principle became available on the market [3]. In recent years, this kind of camera becomes more and more popular and has been widely used in the field of mobile robotics, automated vehicle and interactive entertainment. In particular, some research on obstacle detection [4], pose estimation [5], terrain traversability analysis [6], 3D object reconstruction [7] and simultaneous localization and mapping (SLAM) [8] for mobile robotic applications have been done. At the same time, calibration models of look-up-Tables [9], sinusoidal functions [10] and high-order polynomial [11] are proposed to improve the accuracy of the TOF camera. Using these models can improve the accuracy to some extent, but it also faces the problem of huge amount of data, complex function composition and high-order computation.

In recent years, the SR4500 TOF camera provides a new choice for environment perception in the field of mobile robotics. In order to solve the problem of distance

ISSN: 2005-4297 IJCA

calibration and gray-level image generation of the SR4500, a detailed study is performed in this paper. The main contribution in this context is that the Gaussian Process Regression (GPR) model is used to calibrate the range error. With this model, fewer parameters need to be estimated, and solves the problem of high-order computation and complex function composition when using error models of polynomial or sinusoidal functions. In order to eliminate the error data caused by the object edge, the distance threshold which determined by the range data of the current pixel is selected to remove the mixed pixels. To acquire gray-level images from the amplitude data, the amplitude histogram is used to select the threshold of reflected light, and grayscale images are generated through normalizing the amplitude value with this threshold, by means of this method the generated gray value can be dispersed on the majority of the gray scale range.

The remainder of this paper is organized as follows: the principle of TOF camera is given in Section 2. Section 3 proposes a novel calibration model to improve the range accuracy and the method of removing mixed pixels is also discussed. The algorithm of grayscale image generation is proposed in Section 4. Section 5 shows the experiment of distance calibration, mixed pixel removing and gray-level image generation, and an environment perception experiment is done in an outdoor environment. Finally, a short summary is given in Section 6.

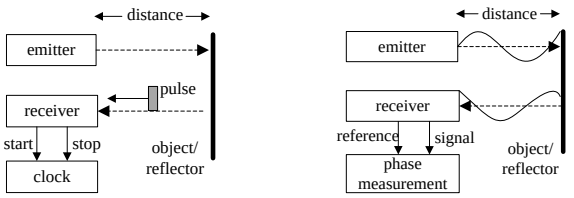
2. TOF Camera Principle

The operation mode of the TOF camera is based on the Time-of-Flight principle. The observed scene is illuminated by modulated infrared light which is reflected by visible objects and gathered in an array of solid-state image sensors. By measuring the travel time of light, the distance at each pixel is determined:

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

where d is the distance between sensor and object; c is the speed of light; t is the time between emitting and receiving.

According to the form of emitted light pulse, there are two different types of Time-of-Flight principle, direct Time-of-Flight principle and indirect Time-of-Flight principle, which are shown in Figure 1. In the first case, the runtime of a travelled light pulse is directly measured using arrays of single-photon avalanche diodes (SPADs) [12]. This distance measurement has high precision and suits for the long-range detection, but In order to reach a distance accuracy of a few millimeters, the clock accuracy has to be as low as a few picoseconds [13].



(a) Direct principle (b) Indirect Principle

Figure 1. Time-Of-Flight Distance Measurement Principle

The other method uses amplitude modulated light and obtains distance information by measuring the phase difference between the reference signal and reflected signal [14]. The emitted light is modulated in amplitude with a sinusoidal modulation. The received

sinusoidal modulated signal is sampled four times in each cycle, at ¼ period phase shifts, *e.g.* 90° phase angle. From the four samples ($c(\tau 0), c(\tau 1), c(\tau 2), c(\tau 3)$), the offset B, the amplitude A and the phase shift φ can be calculated:

$$B = \frac{c(\tau 0) + c(\tau 1) + c(\tau 2) + c(\tau 3)}{4} \tag{2}$$

$$A = \frac{\sqrt{(c(\tau 0) - c(\tau 2))^2 + (c(\tau 1) - c(\tau 3))^2}}{2} \tag{3}$$

$$\varphi = \arctan\left(\frac{c(\tau 0) - c(\tau 2)}{c(\tau 1) - c(\tau 3)}\right) \tag{4}$$

The scheme of the phase shift measurement principle is reported in Figure 2, B is the mean intensity of the received signal and A is its amplitude. The received signal is offset-shifted in intensity with respect to the emitted signal mainly because of the additional background light.

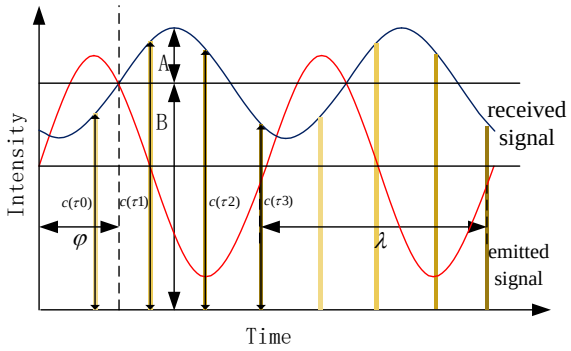


Figure 2. Scheme of the Phase Shift Measurement Principle

The distance d is then derived from the phase shift φ:

$$d = \frac{\lambda}{2} \bullet \frac{\varphi}{2\pi} \tag{5}$$

where λ is the modulation wavelength.

3. Distance Data Processing

3.1. Distance Calibration with GPR

In the last few years, some work on distance calibration has been done and many models are used to adjust the range measurement errors, such as look up Tables model, high-order polynomial model and sinusoidal model. But in this paper, we proposed a novel approach which using the GPR [15,16] model to solve this problem.

The GPR is a new machine learning method by the context of Bayesian theory and the statistical learning theory [17]. It is widely used to solve the high-dimensional, small-sample or nonlinear regression problems. By using a set of training points, a GPR can provide a prediction of the unknown function with an associated uncertainty over a continuous domain [18]. Compared with other models, the GPR model needs no high-order computation and complex function composition and has fewer parameters to be estimated, so it is suitable for this problem.

A GP is determined by its mean function $m(x)$ which is always assumed to be zero and a covariance function $k(x, x')$:

$$f(x) \sim GP(m(x), k(x, x')) \tag{6}$$

where

$$m(x) = E\left[f(x) \right] \tag{7}$$

$$k(x, x') = E\left[(f(x) - m(x))(f(x') - m(x')) \right] \tag{8}$$

For the regression problem:

$$y = f(x) + \varepsilon \tag{9}$$

where f is the function value, y is the observed target value and assuming polluted by the independent identically distributed Gaussian noise $\varepsilon \sim N(0, \sigma_n^2)$.

the prior on the noisy observations becomes

$$\operatorname{cov}(y) = K(X, X) + \sigma_n^2 I \tag{10}$$

where σ_{pq} is a Kronecker delta which is one if p = q and zero otherwise.

Then, the joint distribution of the training outputs y and the test outputs f_* according to the prior is

$$\begin{bmatrix} y \\ f_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(X, X) + \sigma_n^2 I & K(X, x_*) \\ K(x_*, X) & k(x_*, x_*) \end{bmatrix} \right) \tag{11}$$

if there are n training points and n_* testing points then $K(X, x_*)$ denotes the $n \times n_*$ matrix of the covariance evaluated at all pairs of training and testing points, and similarly for the other entries $K(x_*, X)$, $K(X, X)$ and $K(x_*, x_*)$.

The predictive distribution for testing locations x_* is given by

$$\begin{gathered} f_* \mid X, y, x_* \sim \\ N(\bar{f}_*, \operatorname{cov}(f_*)) \end{gathered} \tag{12}$$

where

$$\bar{f}_* = K(x_*, X) K^{-1} y \tag{13}$$

$$\operatorname{cov}(f_*) = K(x_*, x_*) - K(x_*, X) K^{-1} K(X, x_*) \tag{14}$$

with $K = K(X, X) + \sigma_n^2 I$.

Equations (13) and (14) provide the basis of the range error estimation process.

There are numerous covariance functions that can be used to model the spatial variation between the data points, *e.g.* Linear, Exponential, Squared Exponential. A widely used is the squared-exponential kernel given as:

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{1}{2}(x - x')^T \sum^{-1} (x - x')\right) \tag{15}$$

where $\sum = diag(l^2)$, $\sum$ is the length-scale matrix, a measure of how quickly the modeled function changes in the directions x and y; σ_f^2 is the scaling parameter. The set

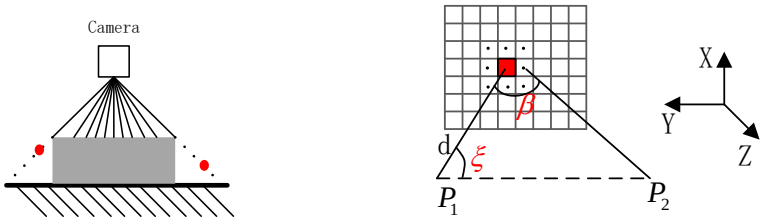
of $\theta = \{\sum, \sigma_f^2, \sigma_n^2\}$ are referred to as the hyperparameters that fully determine the GPR, and are always determined through Maximum Likelihood Estimation.

Because of the range error is the function of one variable about distance, then the covariance function (15) can be rewritten as:

$$k(x, x') = \sigma_f^2 \exp(\frac{-(x - x')^2}{2l^2}) \tag{16}$$

3.2. Mixed Pixels Removing Method

As shown in the Figure 3(a), the range data from the scene contains mixed pixels caused the object edge, which degrades the data acquired on the borders of the object greatly. S. May used angle ξ for thresholding jump edges [1]. In this paper, we use a similar method to remove the mixed pixel.



(a) Mixed pixel (b) Threshold selection

Figure 3. Illustration of Mixed Pixel Removing

Through the distance d of the observation point p_1 at current pixel and the angular resolution β of the camera, we can calculate distance threshold between the neighboring data points p_1 and p_2:

$$threshold = k * 2d * \sin(\beta / 2) \tag{17}$$

where β =0.39°; k is a scale factor.

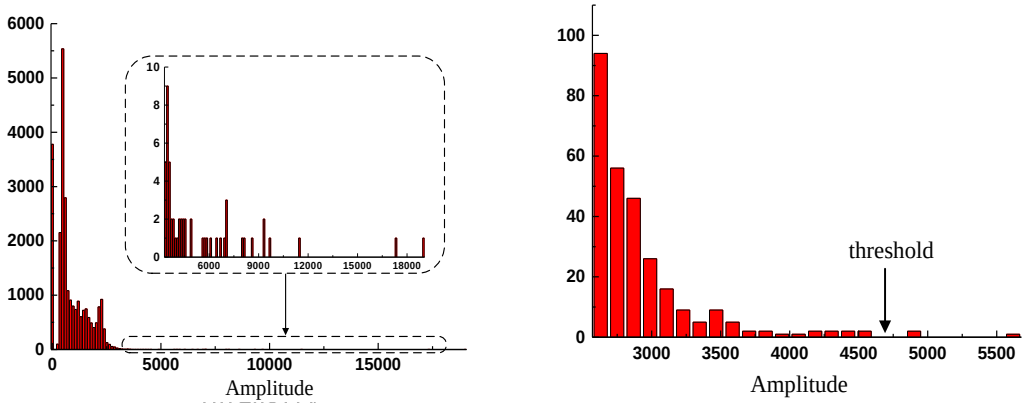
Then, the distance between p_1 and its eight neighborhood points will be calculated. If one of the distances exceeds the threshold, the point p_1 is regarded as mixed pixel.

4 Gray-level Image Generation Algorithm

The SR4500 camera does not generate the color data of the scene but the amplitude information of the reflected signal with the value from 0 to 0xFFFF. In order to improve the availability of the amplitude information of the reflected light, a normalization approach is used to translate the amplitude value to 0-255, and the resulting image is more looking like a conventional gray-level camera image.

In this paper, the amplitude histogram is used to accept the overall situation of the amplitude data, and a threshold is selected out to determine the amplitude upper limit. Through this threshold, the amplitude data is normalized to 0-255.

Figure 4(a) shows the amplitude histogram. We can see that the max amplitude value at this picture is about 18000, but most of the data is less than 5000, and the data which is greater than 5000 are rarely and very scattered. In Consideration of the continuity of the amplitude, the point which's intensity is zero and the number of elements before this point are greater than 98 percent of the total is selected as the break point. For the Figure 4(a), the threshold is shown in the Figure 4(b) and its value is 4542.



(a) Amplitude histogram (b) Threshold selection

Figure 4. Threshold Selection Using Amplitude Histogram

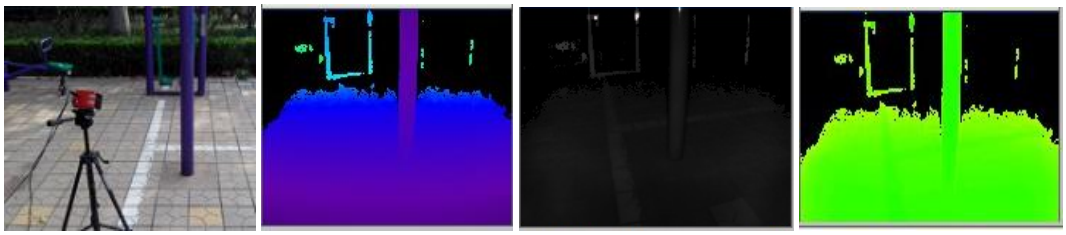
The normalization equation is:

$$P(i, j) = \frac{P(i, j) - \min}{\max - \min} * 255 \tag{18}$$

where $P(i, j)$ is the amplitude value of each pixel; $\max$ is the threshold of the amplitude value and $\min$ is the minimum of all amplitude value

5. Experiment Results

The following sections describe the experiment of distance calibration, mixed pixel removing and gray-level image generation. The SR4500 camera is selected as the experimental device. This camera is the fourth-generation ToF cameras produced by MESA in Switzerland. It is based on the principle of the continuous wave time-of-flight method and can provide range image, amplitude image and confidence image, which are shown in Figure 5.



(a) Scene image (b) Range image (c) Amplitude image (d) Confidence image

Figure 5. Raw Output Images from the SR4500

5.1. Distance Calibration Experiment

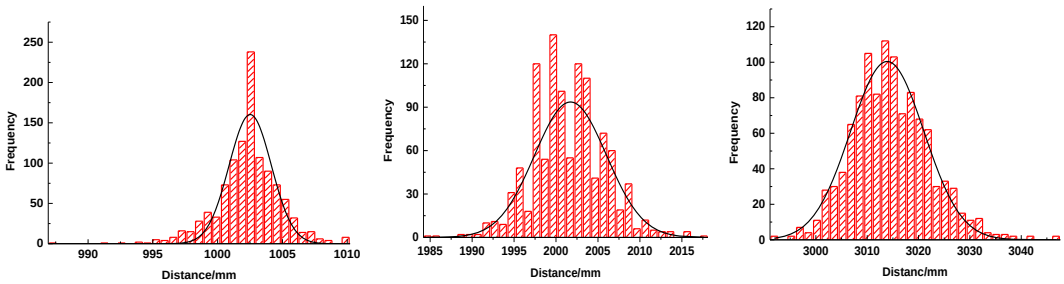
The distance measurement of the SR4500 camera is influenced by many factors, such as integration time, internal temperature and ambient light. Because of the SR4500 combines the optical feedback loop and temperature compensations, so the influence of the integration time is considered only in this paper.

In order to evaluate the systematic distance measurement error of the SR4500 camera, the range error distribution of the center pixel is investigated. The camera is set up on a photographic tripod and paralleled to a vertical white board (Figure 6a). The Figure 6b shows the camera view and the central pixel is at the cross area. The distance between the camera and the board is accurately measured using Bosch DLE40 laser range finder (Figure 6c).

(a) Testing Scene (b) Camera View (c) DLE40

Figure 6. Experiment Setup for Evaluate the Range Errors

The histogram of the 1000 distance measurements measured by the central pixel with an integration time of 10.3ms for the distance of 1006mm, 1983mm and 2993mm is generated. The Figure 7 shows that the measured distance is very close to the approximated value and it follows Gaussian distributions.



(a) 1006mm (b) 1983 mm (c) 2993 mm

Figure 7. Histogram of the 1000 Distance Measurements at Different Distance

In order to study the influence of different integration time, distance measurements with respect to the integration time of 5.3ms, 10.3ms, 15.3ms and 20.3ms at different position are measured. The panel is positioned from 1000mm to 4200mm with five-centimeter interval . Ten consecutive frames are acquired for each panel position, and the mean μ, standard deviation std and range error err are calculated through:

$$\mu = \frac{\sum_{i=1}^{n} d_i}{n} \tag{19}$$

$$std = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \mu)}{n-1}} \tag{20}$$

$$err = \mu - ref \tag{21}$$

where d_i is the distance data delivered by the camera; ref is the distance reference value; n is set to10 in the above equation.

The 95% confidence interval for the range error distribution is shown in Figure 8. The result exhibits that the oscillating error appeared with the increasing of the distance and the range error standard deviation is in inverse proportion with respect to the integration time. On the other hand, though the influence of integration time to the mean of the range error is small, it is greater to the range error standard deviation. Considering the stability of the range data, the integration time which is greater than 10.0ms is suggested to be selected in the future application.

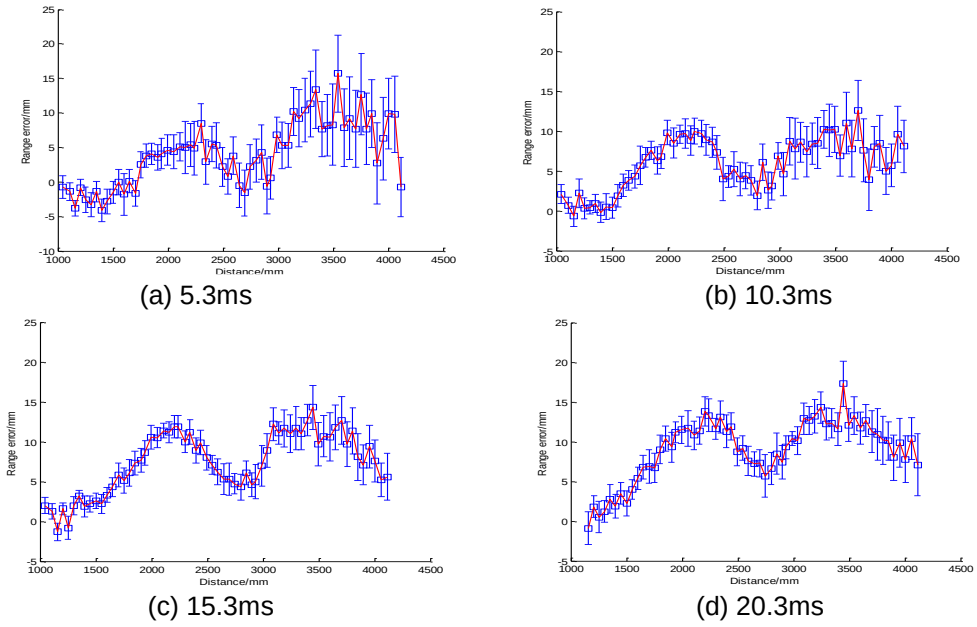


Figure 8. Range Error of the Center Pixel at Different Integration Time

From Figure 8 we can see that the range error is not the same at different integration time but the difference between each other is small, so a set of range error is selected from the four as training points to predicate the error at arbitrary distance, and the range data at the other integration time is calibrated by this group of predicted value. The selection is realized through calculating the sum of squared difference with each other, and the Figure 8(a) is excluded because of the big fluctuation. The result is shown in Figure 9. From this figure we can see that the range error with the integration time of 15.3ms is less than the other so it is selected as the training points.

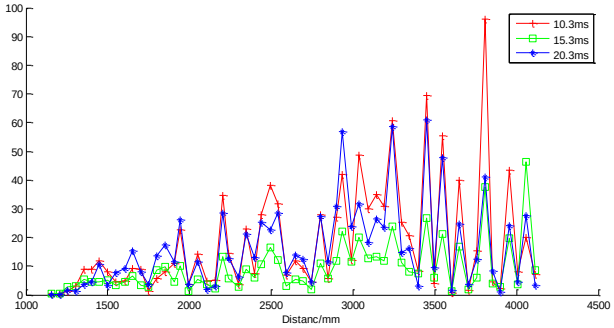
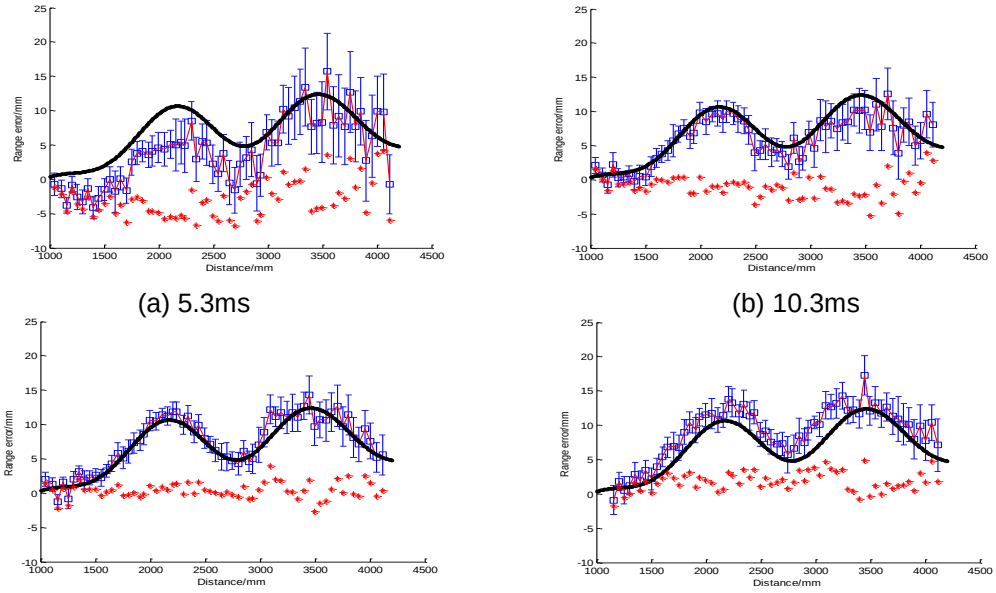


Figure 9. Training Points Selection

Then, the equation (13) and equation (16) are used to predict the range error from 1000mm~4200mm. The result is shown in Figure 10, the black line is the estimated value of range error and the red points represent the calibrated range error. As can be seen from this Figure, after applying the GPR error model the range error goes down from 15mm to 5mm in the 1000~4200 mm distance measurement, and it is can meet the requirements of perception in the field of robotics.



(a) 5.3ms (b) 10.3ms

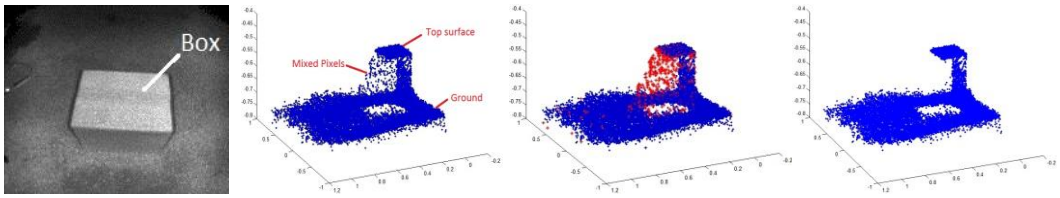
(c) 15.3ms (d) 20.3ms

Figure 10. Calibration Result at Different Integration Time

5.2. Mixed Pixels Removing Experiment

This experiment illustrates the performance of the mixed pixels removing algorithm. The level ground with a box placed there is selected as the experimental subject, which is shown in Figure 11(a).

Figure 11(b) is the raw 3D scene acquired by the SR4500, we can see that there are many scatter plots between the top surface of the box and the ground. In order to remove these points, the equation (17) is used to select the mixed pixel. Figure 11(c) shows the result and the mixed pixels are shown with red color. Figure 11(d) is the 3D scene after removing mixed pixels. We can see that though this method is convenient to realize, the mixed pixels can be removed well.



(a) Scene Image (b) 3D Scene (c) Pixels Selection (d) Pixels Removing

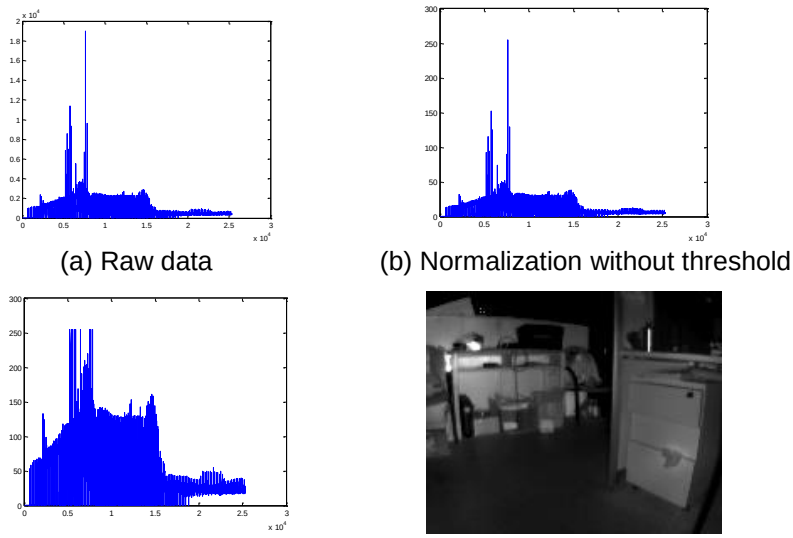
Figure 11. Mixed Pixels Removing

5.3. Gray-level Image Generation Experiment

An amplitude image captured by the SR4500 camera is used to verify the gray-level image generation algorithm, and the result is shown in Figure 12.

Figure 12(a) is the raw amplitude data and the maximum is about 19000; Figure 12(b) is the normalization result of using the maximum of amplitude value, this method can acquire the grayscale data but the majority of the value is less than 50; Figure 12(c) is the normalization result of using the threshold and most of the data is greater than 100; Figure 12(d) is the gray-level image generated with the proposed algorithm. From this experiment we can see that by means of proposed method, the amplitude value can be

dispersed on the majority of the gray scale range and increases the resolution of the gray value.



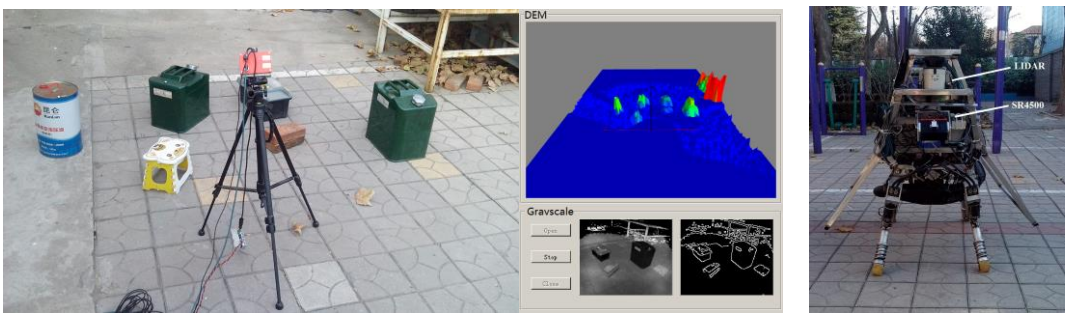
(a) Raw data (b) Normalization without threshold

(c) Normalization with threshold (d) Gray-level image

Figure 12. Normalization of the Amplitude Value

5.4. Outdoor Experiment

Based on the analysis above, a perception experiment in an outdoor environment is done, which is shown in Figure 13(a). The Digital Elevation Map (DEM) is generated according to the range data. From this experiment we can see that the precise 3D environment map and grayscale images can be obtained in the outdoor environment and the grayscale image can adapt to high background light levels. At the same time, we have installed this camera in the quadruped robot, and researches of mapping and obstacle detection are ongoing.



(a) Environment perception experiment (b) Application

Figure 13. Outdoor Experiment

6. Conclusion and Future Work

ToF cameras can provide both range images and amplitude images at video frame rates without scanning and from just one point of view, so it is well suited for the environmental perception under complex condition. In this paper, some tests relative to the new and more powerful SR4500 camera have been reported, and three novel approaches are proposed to solve the problem of range error calibration, mixed pixels removing and gray-level image generation.

1) First, compared with distance error model using polynomial or sinusoidal functions, the GPR model is used in this paper. With this model, fewer parameters need to be estimated and the high-order computation and complex function composition are avoided. After been calibrated, the range error goes down from 15mm to 5mm.

2) Second, In order to eliminate the error data caused by the object edge, the distance threshold which determined by the range data of the current pixel is selected to remove the mixed pixel. This method is convenient to realize and can remove the mixed pixel well.

3) Third, the amplitude histogram is used to select the threshold of reflected light, and gray-level images are generated through normalizing the reflected light with this threshold. By means of this method the generated gray value can be dispersed on the majority of the gray scale range.

These three approaches enable the usage of this camera, and based on the proposed methods, the work on environment information collection for mobile robot is ongoing. Besides that, further studies should be done in order to fully use the sensors abilities and exploit its potentialities, such as obstacle detection, terrain information recognition, pose estimation and simultaneous localization and mapping.

Acknowledgements

Supported by the National Natural Science Foundation of China (61233014, 61305130 and 61503153),the National High Technology Research and Development Program of China (2015AA042201) the Shandong Provincial Natural Science Foundation (ZR2013FQ003 and ZR2013EEM027), and China Postdoctoral Science Foundation (2013M541912).

References

[1] S. May, S. Fuchs, D. Droeschel, D. Holz and A. Nuchter, "Robust 3D-Mapping with Time-of-Flight Cameras", Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, USA, **(2009)** October 11-15, pp.1673-1678.

[2] F. Chiabrando, D. Piatti and F. Rinaudo, "SR-4000 ToF camera: further experimental tests and first applications to metric surveys", International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Newcastle, UK, Vol. XXXVIII Part 5, **(2010)**, pp. 149-154.

[3] T. Oggier, M. Lehmann, R. Kaufmannn, M. Schweizer, M. Richter, P. Metzler, G. Lang and F. Lustenberger, N. Blanc, "An all-solid-state optical range camera for 3D-real-time imaging with sub-centimeter depth-resolution (SwissRanger)", Proceedings of SPIE, St. Etienne, France , Vol. 5249, **(2003)**, pp. 534-545.

[4] S. Hussmann, T. Ringbeck and B. Hagebeuker, "A performance review of 3D TOF vision systems in comparison to stereo vision systems", in Stereo Vision (Online book publication), Edited Asim Bhatti, I-Tech Education and Publishing, Vienna **(2008)**, ch. 7, pp. 103-120.

[5] S. Hong, C. Ye, M. Bruch and R. Halterman, "Performance evaluation of a Pose Estimation method based on the SwissRanger SR4000", Proceedings of 2012 IEEE International Conference on Mechatronics and Automation, **(2012),** Chengdu, China, August 5-8, pp. 499-504.

[6] I. Bogoslavskyi, O. Vysotska, J. Serafin, G. Grisetti and C. Stachniss, "Efficient Traversability Analysis for Mobile Robots using the Kinect Sensor", Proceedings of 2013 European Conference on Mobile Robots (ECMR), Barcelona, Spain, **(2013)** September 25-27, pp. 158-163.

[7] B. Dellen, G. Alenyà, S. Foix and C. Torras, "3D object reconstruction from Swissranger sensor data using a spring-mass model", Proceedings of the 4th Internal Conference Computer Vision Theory and Application-Cetto, Lisbon, Portugal, **(2009)**, Vol. 2 pp. 368-372.

[8] K. Pirker, M. Ruther, H. Bischof, G. Schweighofer and H. Mayer, "An omnidirectional Time-of-Flight camera and its application to indoor SLAM", Proceedings of the 11 Internal Conference on Control Automation Robotics & Vision, Singapore, **(2010)** December 7-10, pp. 988-993.

[9] T. Kahlmann, H. Remondino and H. Ingensand, "Calibration for increased accuracy of the range imaging camera SwissRanger", International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Dresden, Vol. XXXVI, Part 5, **(2006)**, pp. 136-141.

[10] F Chiabrando, R. Chiabrando and D. Piatti, F. Rinaudo, "Sensors for 3D Imaging: Metric Evaluation and Calibration of a CCD/CMOS Time-of-Flight Camera", Sensors, **(2009)**, pp. 10080-10096.

[11] Y. M. Kim, D. Chan, C. Theobalt and S. Thrun, "Design and Calibration of a Multi-view TOF Sensor

Fusion System",. Proceedings of IEEE CVPR Workshop on Time-of-flight Computer Vision, **(2008)**.
[12] M. Lindner, I. Schiller, A. Kolb and R. Koch, "Time-Of-Flight sensor calibration for accurate range sensing", Computer Vision and Image Understanding, Vol. 114, **(2010)**, pp. 1318-1328.
[13] M. A Albota, R. M Heinrichs, D. G Kocher, D. G Fouche, B. E Player, M. E Obrien, G. F Aull, J. J Zayhowski, J. Mooney, B. C Willard and R. R Carlson, "Three-dimensional imaging laser radar with a photon-counting avalanche photodiode array and microchip laser", Applied Optics, Vol. 41, Iss. 36, **(2002)**, pp. 7677-7678.
[14]
[15] R. Lange, Time-of-Flight range imaging with a custom solid-state image sensor. Proceedings of SPIE, Munich, Germany, **(1999)**, Vol. 3823, pp. 180-191.
[16] C. E. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning. MIT Press, **(2006)**.
[17] Z. He, G. Liu, X. Zhao and M. Wang, "Overview of Gaussian process regression", Control and Decision, **(2013)**, Vol. 28, No. 8, pp. 1121-1129.
[18] S.i Vasudevan, F. Ramos, E. Nettleton, H. Durrant-Whyt and A. Blair, "Gaussian Process Modeling of Large Scale Terrain."
[19] J. R. Souza, R. Marchant, L. Ott, D. F. Wolf and F. Ramos, "Bayesian Optimisation for Active Perception and Smooth Navigation", Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, **(2014)** May 31- June 7, pp. 4081-4087.

