Creating a Correction Table for the Purpose of Improving the Accuracy of the RGB-D Sensor

Shinya Fujino*, Masaki Ishii*, Shuntaro Sato*

*a Akita Prefectural University, 84-4 Tsuchiya-Ebinokuchi, Yurihonjo City 015-0055, Japan

*Corresponding Author: ishii@akita-pu.ac.jp

Abstract

In recent years, the RGB-D sensor has attracted attention in the field of 3D mapping. This sensor is equipped with a camera, an infrared camera, and an infrared projector. An RGB-D sensor is able to acquire both depth and color images. However, because it is not a sensor for the purpose of accurate measurement, the depth image is distorted.

Our work aimed to use an RGB-D sensor to enable a robot to move around indoors. Therefore, we developed a calibration technique for the RGB-D sensor with the purpose of building accurate 3D maps\(^{1(2)}\). However, we encountered two problems because the distortion of the depth image was larger than intended. First, the accuracy we were able to achieve was unacceptable. Second, a discontinuous point was found to exist in the correction formula.

In this research, we derived a correction formula by starting with a linear function and working upwards until a septic function was obtained to solve the problem. We carried out a polynomial approximation using a least-squares method to derive the correction formula. In addition, the correction table was created by selecting an optimum correction formula for each pixel. The corrected depth image was used to construct 3D maps and the accuracy of these maps is discussed.

Keywords: RGB-D, calibration, point cloud, 3D SLAM.

1. Introduction

Currently, robots are used in various situations, because of the great expectations that robot technology would enrich society. The ability to generate a 3D map is an important part of robotic technologies that are required for robots intended to provide a service in living spaces. An environmental map composed of geometrical features such as the position of an obstacle is necessary for the robot to move autonomously. Thus, robots require an autonomous 3D map-building technique because of the difficulties associated with building a pre-environment map for a variety of environments, such as human living spaces.

In previous studies, several calibration methods were studied. For example, there is a method to approximate the cylindrical model\(^ {3}\). On the other hand, there is also a method of correcting to derive a correction equation using short-range data\(^ {4}\). However, both methods are difficult to correct long range data.

In this research, we derived a correction formula by starting from a linear function and working upwards until we obtained a septic function to solve the problem. We conducted a polynomial approximation using a least-squares method to derive the correction formula. In addition, we created the correction table by selecting an optimum correction formula for each pixel.

2. Proposed Method

2.1 Calibration by the Polynomial Approximation

In this study, we use the Xtion PRO LIVE as RGB-D sensor. In previous studies, we derived the linear function for correction by using a least-squares method for each pixel for
the purpose of reducing the distortion of the depth image obtained. The results show that it is possible to accurately correct the point cloud data obtained from the RGB-D sensor by using two kinds of correction formulas, i.e., for both short and long range. However, in this regard two problems exist because the distortion of the depth image was far larger than intended. The first problem is that the depth image was not accurate, whereas the second problem is that a discontinuous point exists in the correction formula.

In this work, we derived a correction formula by starting from a linear function and working upwards until we obtained a septic function to solve the problem. We carried out a polynomial approximation using a least-squares method to derive the correction formula. In addition, we created the correction table by selecting an optimum correction formula for each pixel.

The ratio of the actual distance D and the depth Z was approximated by the n-th order function using the calculated least-squares method for each distance. We derived a correction equation by calculating each coefficient.

\[ \frac{D}{Z} = a_n Z^n + a_{n-1} Z^{n-1} + \cdots + a_2 Z + a_0 \]  
(1)

\[ Z_t = (a_n Z^n + a_{n-1} Z^{n-1} + \cdots + a_2 Z + a_0) \times Z \]  
(2)

After calculating \( Z_t \), corrected coordinates \( X_t \) and \( Y_t \) were calculated using the angles \( \alpha \) and \( \beta \) for the X- and Y-axes, respectively.

\[ X_t = Z_t \tan \alpha \]  
(3)

\[ Y_t = Z_t \tan \beta \]  
(4)

2.2 Correction Table

Distortion of the depth image sensor is complicated because the factory calibration error, lens distortion, and noise are combined. In our previous study, we derived a corrected linear function for each pixel. However, it is necessary to approximate each pixel since the tendency of strain is different for each pixel.

In the work presented in this paper, we created the correction table by selecting an optimal correction formula.

The results show that it is possible to accurately correct the point cloud data obtained from the Xtion sensor. The depth image was obtained in the Honjo campus gym of the Akita Prefectural University. Figure 2 shows the data acquisition environment. In this study, the depth image is obtained by increasing the distance between the Xtion sensor and the wall by a constant distance.

At that time, the floor is captured in the lower portion of the point cloud data at an actual distance of 3.0 [m] or more of acquired data. Therefore, the data was acquired by installing the Xtion upside down. Figure 3 shows the depth image at the point \( D = 5.0 \) [m] as an example of the acquired data. The results of the analysis of the plane depth image revealed a lack of data in the image of the actual distance 0.5 \( \sim 0.7 \) [m].

At an actual distance of 8 [m] or more, the acquisition of stable data for images is difficult. Therefore, we used an image of the real distance 0.8 \( \sim 8.0 \) [m]. Incidentally, the origin is the position of the Xtion and the X-, Y-, and Z-axes represent the horizontal, vertical, and depth directions, respectively. The coordinates of each pixel of the point cloud data are expressed in the 3D coordinate system (X, Y, Z). \( D \) [m] denotes the actual distance to the wall, and \( Z \) [m] denotes the resulting depth from the Xtion.

3. Calibration Experiment

3.1 How to Obtain the Depth Image

The correction formula of the calibration method in this study is derived by using the plane depth image obtained by the Xtion sensor. The depth image was obtained in the Honjo campus gym of the Akita Prefectural University. Figure 2 shows the data acquisition environment. In this study, the depth image is obtained by increasing the distance between the Xtion sensor and the wall by a constant distance.

We derived a correction equation using a plane depth image, and corrected the equation for depth images. Figure 4 shows the average error of each correction equation. The figure shows that the average error is reduced to an
increasing degree, although the average error of the cubic function did not improve and the average error of the sextic and septic functions increased. Table 1 shows the number of corresponding pixels in the case of focusing on the coefficient of determination of each correction formula. The number of pixels with a large coefficient of determination increased as the degree of the function increased.

However, the number of pixels with a large coefficient of determination was greatly reduced in the fourth order. Even greater pixel of the coefficient of determination orders rise also increased. Therefore, although the coefficient of determination improved, it shows a behavior curve is not assumed, and that the average error is increasing. This result indicated that, when considering the calculation cost and average error, the correction equation is considered to be sufficient when the cubic equation is reached. Therefore, the correction table is generated using three correction equations from a linear to a cubic expression.

### 3.3 Result of Creating the Correction Table

The coefficients of determination of the three correction formulas were compared at each pixel based on the knowledge obtained in Section 3.2. The correction table was created by selecting a correction formula with the highest coefficient of determination. Figure 5 shows the correction table that was created. Table 2 provides the proportion of correction formulas that exist in the correction table. The results in Figure 5 and Table 2 indicate that the cubic expression is selected for most pixels and that the linear and quadratic equations were selected by a small number of pixels.

Therefore, calibration of the RGB-D sensor is required for each pixel. Next, Figure 6 shows the average error of each correction equation when the use of a correction table is included. The average error of the correction table is consistent with being substantially cubic. However, the error improved for longer distances of 6.0 [m] or more. As a result, the proposed method is considered useful for correcting large variations in long-range data.

### 4. 3D Mapping Experiment

Previous studies proposed a registration method (SIFT+PFHRGB method) using a depth image based on the initial posture estimated by the registration method (ICP) using a color image. Figure 7 illustrates the processing procedure. In this study, the same method was used to construct the map. This experiment was conducted at Akita
Prefectural University. Figure 8 shows a schematic view of the experimental environment. The arrows in the figure indicate the data acquisition route. This experiment focused on a small distortion of the “straight” and “corner” paths of the larger experiment in our prior study. The data were acquired using an RGB-D sensor that was mounted on a Roomba (iRobot Corp.). Data were acquired at a speed of 0.3 [m/s], rotation of 0.12 [rad/s], and an acquisition interval of 0.7 [fps]. The best results were obtained for the correction that was performed using a correction table. Figure 9 shows the resulting environmental map that was built before and after correction. A long the straight path, the map has reduced distortion and overall contraction. In addition, the overall distortion of the corner has also been alleviated. These results confirm that the proposed calibration method for the RGB-D sensor is useful for map construction.

5. Conclusions

This paper proposes the use of a correction equation obtained by polynomial approximation to create a correction table. As a result, we solved the problem of discontinuities that was experienced by the proposed approach when continuous functions were used to perform the approximation. Further, the accuracy was improved by the selection of an optimal correction formula for each pixel. The usefulness of the proposed method was confirmed by the reduction in map distortion that was achieved by performing the correction using the correction table. This correction can be expected to be applied in a variety of fields because it is also available in non-construction map.

However, the constructed map we used in this work was a local map, and only exhibited a small difference between each of the correction formulas and the correction table. Therefore, the constructed maps did not confirm the obvious usefulness of the correction table. In the future, we plan to conduct experiments targeting a single with depth in order to quantify the error between the real environment and depth. Furthermore, we are planning to evaluate the usefulness of the proposed method by building a map in a wide-area environment.

References