TRANSFORMATION OF DEPTH MAPS PRODUCED BY TOF CAMERAS

Maciej Kurc, Krzysztof Wegner, Marek Domański
Chair of Multimedia Telecommunications and Microelectronics,
Poznań University of Technology,
Polanka 3, 61-131 Poznań, Poland,
email: [mkurc, kwegner]@multimedia.edu.pl
email: domanski@et.put.poznan.pl

***Abstract* — In this paper we propose a depth map transformation to be used with Time-of-Flight (ToF) depth cameras. ToF cameras measure the distance along light rays while depth map representation uses the Cartesian, z-distance. Necessary correction procedure is usually provided by the camera manufacturer, however the parameters used and algorithm details remain undisclosed. We propose a simple depth map transformation based on a geometrical relations, defined by the pinhole camera model, and independently estimated camera intrinsic parameters. The transformation was evaluated using a series of ToF measurements of a flat surface. The proposed method was compared with the post- processing method implemented by the manufacturer.**

Keywords — ToF; Depth Camera; Depth Maps; Geometric Correction; Depth Transformation

I. INTRODUCTION

A Time-of-Flight (ToF) camera is a device which is capable of measuring distance between each pixel and it's correspondent in the 3D world. The distance measurement is performed by emitting a light wave which is reflected by objects in the scene and returns to the camera sensor. By taking the light wave travel time for each pixel of the image and converting its value to distance, a depth map can be obtained. The depth map has a form of an image, where each pixel value represents the distance between the camera and a corresponding point in 3D world [1].

The ToF camera measures the distance directly along a light ray propagation path which corresponds to the spherical distance. However, depth maps represent the Cartesian z-distance measured along the camera optical axis direction. In order to derive Cartesian distance from spherical distance a transformation is required. The situation is shown in Figure 1.

A required transformation function is usually provided by the camera manufacturer. The method transforms distance measurements using built-in set of camera intrinsic parameters. The lens distortion correction is performed as well. Unfortunately, the parameters used are not known to the end user of the camera. This poses a problem, when transformed depth maps are to be used for further processing for example in FTV television systems [2].

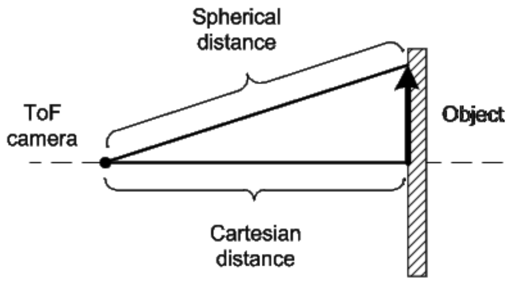

Fig. 1. The spherical and the Cartesian distance.

There are several ToF measurement calibration and transformation methods known from the literature. Kahlmann et al. investigated the ToF depth measurement down to the camera sensor construction [3]. The authors propose a method of distance measurement correction and distance fixed pattern noise estimation using an independent distance measuring device and flat object of reference (white wall). A similar issues are described in [4]. The authors propose the correction of a distance error measurement using a B Spline distortion model both for the spherical and the Cartesian distance. They claim, that both methods leads to the same results, however the accuracy of the distance transformation between coordinate systems itself is not considered.

The distance transformation method, proposed by the authors, provides an alternative to the camera manufacturer's method. It is based on simple geometrical relations of the pinhole camera model. Our proposition does not require any additional camera calibration procedures but the estimation of camera intrinsic parameters. As the algorithm is based on the pinhole model, the lens distortion needs to be removed prior to the processing. The proposition allows to translate the distance with even greater accuracy than the method provided by the camera manufacturer.

II. PROBLEM DEFINITION

A ToF camera provides a distance map in which each measurement is taken along the path between the pixel on the camera sensor and a corresponding point in the 3D world. The way that the distance is measured does not correspond to the depth map representation used by most depth processing algorithms. Visualization of reconstructed wall from distance measured by our ToF camera at approximately 20 degree angle and distance 2.5 m is presented in the Figure 2. Please notice, that lens distortion was already removed. The reconstructed surface is bent due to spherical distance measurement.

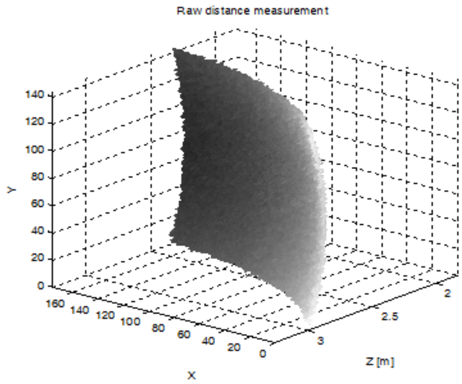

Fig. 2. 3D reconstruction of distance data captured by the ToF camera. X and Y axes are scaled in pixels while Z axis (distance) is scaled in meters.

The data set does not resemble a flat surface. The corners of the observed plane section are located further away from the camera than the center. The Figure 3 illustrates geometrical relations used in pinhole camera model.

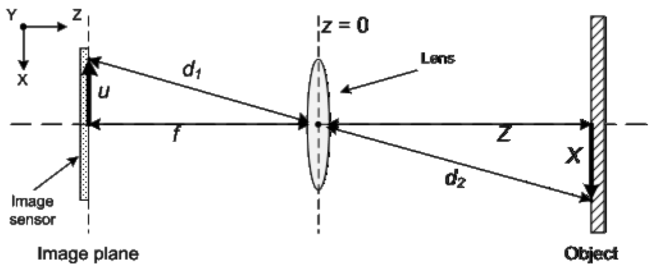

Fig. 3. Illustration of geometrical relations present in the pinhole camera model.

The measured distance value is equal to $d=d_1+d_2$, and the distance, used in depth map representation, corresponds to the value of z. In order to make measurements compliant with the common depth map definition a distance transformation need to be performed. The goal is to find the value of z while knowing the value of d.

III. Proposed Solution

The proposed distance transformation is aimed at transforming a measured distance value to a corresponding depth value (distance measured along the optical axis of the camera). The key advantage of this method is that it relies only on intrinsic parameters of the camera. The lens distortion correction needs to be performed prior to application of the proposed algorithm. The pinhole camera model itself does not represent image distortions.

By using geometrical relations, that can be found in triangles formed by z, x, d_2 and f, u, d_1 line segments, a final distance transformation equation can be derived:

$$z = d \cdot \frac{f_x}{\sqrt[2]{u^2 + f_x^2}} \cdot \frac{f_y}{\sqrt[2]{v^2 + f_y^2}} - f' \qquad (1)$$

where d is the measured spherical distance, z is the transformed Cartesian distance, u and v are the image spatial coordinates (the origin is defined by the optical axis position), f_x and f_y are focal lengths expressed in horizontal and vertical sampling periods and f' represents the distance between the lens and the image sensor.

Focal lengths and principal point position need to be estimated using a camera intrinsic parameter estimation algorithm. Focal lengths f_x and f_y are expressed in horizontal and vertical sampling periods of the image sensor. The value of f' represents the physical focal length.

IV. Experiments

Experiments conducted by the authors are aimed at determining the accuracy of the proposed distance measurement transformation method with respect to the method provided by the ToF camera manufacturer. There are several components that influence the correction accuracy, such as accuracy of the intrinsic camera parameters estimation. The most distorted regions of the depth map are in the corners of its image. In the corners the difference between the actually measured distance and the distance measured along the optical axis is the greatest. The correction inaccuracy is supposed to have its maximum value also in image corners.

A. Methodology

The proposed method requires the ToF camera's intrinsic parameters to be known. Also the lens geometrical distortion needs to be estimated and removed from the acquired data prior to the proposed distance transformation. So we have estimated intrinsic parameters and lens distortion parameters using the Zhang's algorithm [5] and all acquired data was processed first in order to remove lens geometrical distortion.

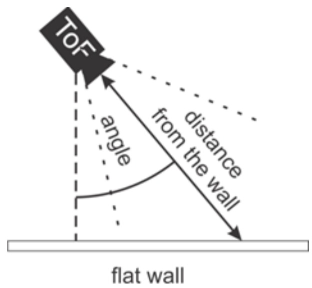

Fig. 4. Experimental setup used in evaluation of the proposed distance distortion correction algorithm.

In order to verify our distance transformation algorithm we have captured depth maps of several large flat surfaces. The Time-of-Flight camera that we use is MESA Imaging SR4000 [6]. The camera is capable of measuring distance at resolution of 176x144 pixels. The modulation frequency was set to 15 MHz which yields in maximum distance range of 10m. The camera was placed in front of a flat wall so that only the wall was visible (Fig. 4). We have captured the depth maps of the wall seen from four distances: 1.5, 2.0, 2.5, 3.0 meters and at multiple viewing angles ranging from 0 to 50 degrees in order to prove that our distance distortion correction algorithm works independently of the distance from the wall and the viewing angle. Each time 50 consecutive frames were captured and averaged to reduce temporal noise.

As the wall was flat, the measured data should represent a continuous flat (planar) surface. We have fitted 3D plane model (Eq. 2) into the measured data both before and after transformation. To fit the model we have used least mean square approach.

$$z = a \cdot u + b \cdot v + c \quad (2)$$

The value of z represents measured depth value (distance from the camera), u and v are coordinates of each pixel of the captured depth map and a, b, c are the plane equation coefficients. Once the plane model is fitted to the data, we have calculated its fitting error by taking the root mean square (RMS) difference between the plane coordinate z and the measured distance data.

We have fit the plane model into every captured depth map for each distance, angle pair. The fit was performed on original data, transformed data using the camera manufacturer algorithm and transformed data using our transformation algorithm.

B. Experimental Results

The figure 5 shows the transformed version of the data set shown in the figure 2. The transformation was performed using our proposed method. The visualization now shows a flat, planar surface.

The figures 6, 7 and 8 show RMS error of the fitted plane model prior to the transformation. For larger distances and angles it was impossible to make the camera see only the wall as the built-in lens provides wide viewing angle. The figure 6 shows data for original data, figure 7 for data processed by camera manufacturer algorithm and 8 for data translated by the proposed algorithm.

The distance transformation procedure, provided by the camera manufacturer, allows to fully counteract the effect of spherical distance measurement. The RMS error stays on the same level independently from the distance and angle change.

The method proposed by the authors performs even better than the one provided by the manufacturer. For small angle and distance, the RMS error value is much lower. However, for angles greater than about 25 degrees the transformation quality begins to deteriorate. One possible explanation is the inaccuracy of estimated lens distortion coefficients. Such inaccuracy often manifests itself on image border. For large camera rotation, the information on image border becomes important as larger parts of the plane surface are projected onto it. Figure 9 shows the plane fitting RMS error difference between the proposed transformation method and the manufacturer method. Negative values favors the proposed algorithm while positive values - the manufacturers algorithm.

Figures 10 and 11 shows percentage of plane model fitting error reduction for the manufacturer and the proposed method. The reduction is shown with respect to the original data set.

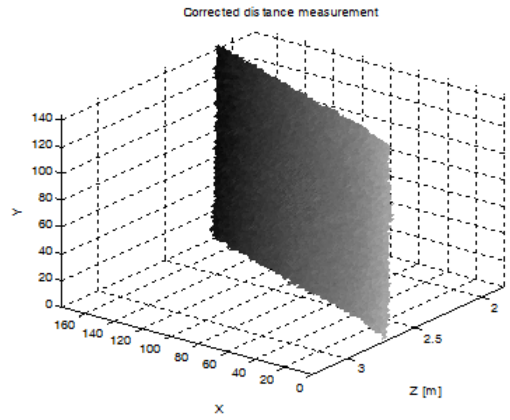

Fig. 5. Wall surface reconstruction after distance distortion correction. X and Y axis are scaled in pixels while distance is scaled in meters.

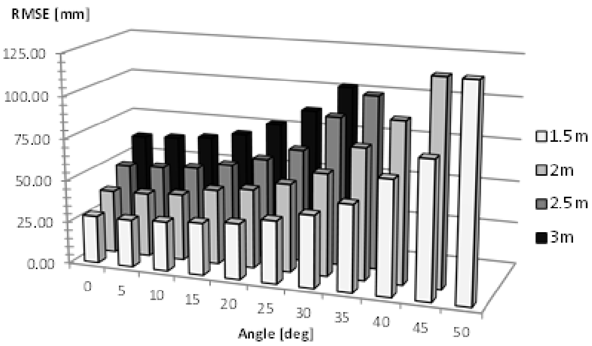

Fig. 6. Diagram of plane fitting RMS error for the original data set.

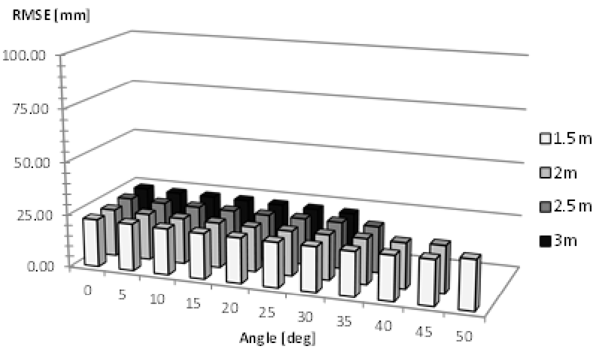

Fig. 7. Diagram of plane fitting RMS error for the data set processed by camera manufacturer distance translation method.

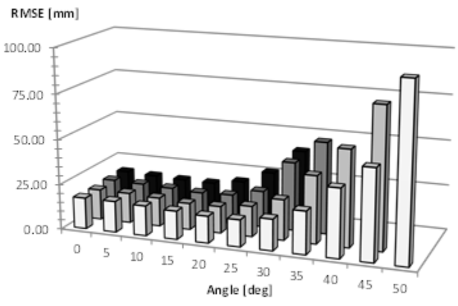

Fig. 8. Diagram of plane fitting RMS error for the data set processed by the proposed translation algorithm.

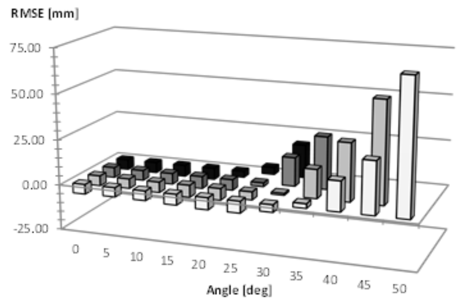

Fig. 9. Diagram of the plane fitting RMS error difference between the proposed method and the camera manufacturer method. Negative values favor the proposed algorithm while positive the manufacturer algorithm.

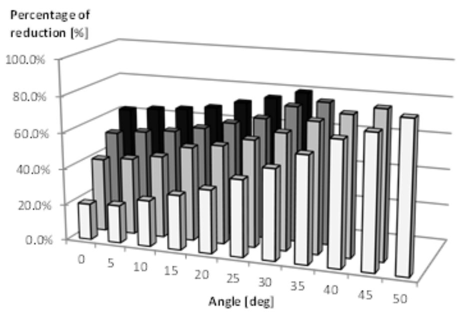

Fig. 10. Diagram of percentage of RMS error reduction for the manufacturer method.

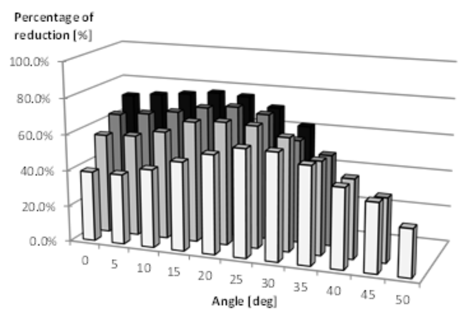

Fig. 11. Diagram of percentage of RMS error reduction for the proposed method.

Figure 12 shows the percentage error reduction difference. For the method, provided by the camera manufacturer, an error reduction chart shape resembles the RMS error shape for the original data. For more distorted images, more of the distortion is removed independently from the absolute distortion amount. On the other hand, the proposed method allows more accurate transformation for small camera rotation angles (below 30-35 degrees).

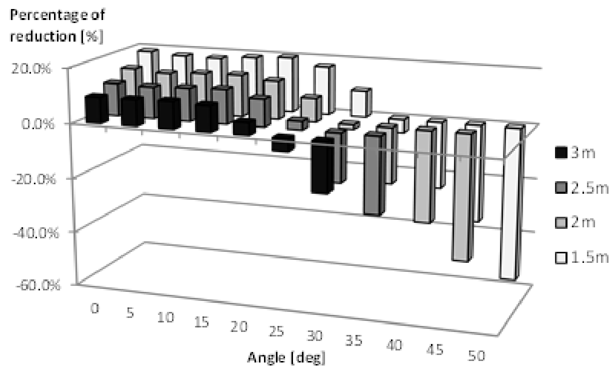

Fig. 12. Diagram of percentage of RMS error reduction difference between the proposed method and the manufacturer method. Positive values favor the proposed algorithm while negative values the manufacturer algorithm.

V. Conclusions

In this paper we propose the method of distance transformation for ToF cameras that transforms the acquired distance map to commonly used depth map representation. The method is based on simple geometric relations and require only camera's intrinsic parameters. The proposed method is an alternative to the similar method provided by the ToF camera manufacturer which details remain undisclosed. In most cases, the proposed method yields in better results than the technique provided by the camera manufacturer which was proven by the experiment conducted by the authors.

Acknowledgments

This work has been supported by the public founds as a DS research project.

References

[1] O.Choi, R.Horaud, "Time-of-Flight Cameras – Principles, Methods and Applications", *Springer* 2013.

[2] M.Tanimoto, "Overview of FTV (free-viewpoint television)", IEEE International Conference on Multimedia and Expo (ICME) 2009, pp. 1552-1553, June 2009.

[3] T. Kahlmann, F. Remondino, H. Ingensand, "Calibration for Increased Accuracy of the Range Imaging Camera Swissranger™", ISPRS Commision V Symposium – Image Engineering and Vision Metrology, Dresden, Germany, Sept. 2006.

[4] M.Linder, A.Kolb, "Lateral and Depth Calibration of PMD-Distance Sensors", Proceedings of the Second international conference on Advances in Visual Computing (ISVC), vol 2.,pp. 524-533, 2006.

[5] Z.Zhang, "A Flexible New Technique for Camera Calibration", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.22, no.11, pp.1330,1334, Nov 2000.

[6] MESA Imaging SR4000 Time-Of-Flight camera, http://www.mesa-imaging.ch