

A Calibration-and-Error Correction Method for Improved Texel (Fused Ladar/Digital Camera) Images

Scott E. Budge

Center for Advanced Imaging Ladar, Utah State University,
Logan, UT 84322-4170, (435) 797-3433

ABSTRACT

The fusion of imaging ladar information and digital imagery results in 2.5-D surfaces covered with texture information. Called “texel images,” these datasets, when taken from different viewpoints, can be combined to create 3-D images of buildings, vehicles, or other objects. These 3-D images can then be further processed for automatic target recognition, or viewed in a 3-D viewer for tactical planning purposes.

This paper presents a procedure for calibration, error correction, and fusing of ladar and digital camera information from a single hand-held sensor to create accurate texel images. A brief description of a prototype sensor is given, along with calibration technique used with the sensor, which is applicable to other imaging ladar/digital image sensor systems. The method combines systematic error correction of the ladar data, correction for lens distortion of the digital camera image, and fusion of the ladar to the camera data in a single process. The result is a texel image acquired directly from the sensor. Examples of the resulting images, with improvements from the proposed algorithm, are presented.

Keywords: lidar, ladar, calibration, sensor fusion

1. INTRODUCTION

Interest in the generation of 3D imagery has been increasing in recent years, with the introduction of applications such as 3DTV, games involving avatars, virtual reality, and historic site documentation. In addition, the capability of creating 3D images is of defence and security interest in automatic target recognition (ATR) and tactical planning. The ability to generate these images directly from sensors and without significant post-processing will enable new and exciting real-time 3D image applications.

There has been much interest in combining information obtained from multiple sensors at a single location in order to create 2.5-D surfaces covered with texture information. The most straightforward method is to use a pair of cameras in a stereo configuration. This approach has the drawback that depth information can be difficult to determine inaccurately in areas of the images with low detail or low contrast. One solution is to add a depth sensor (such as a ladar) to the stereo pair and combine the measurements to reduce the error.^{1,2}

A lower-cost approach is to fuse the data from a single camera and ladar to create a 2.5-D image. If the 3D information is fused to the digital image at the pixel level by the sensor, the image can be thought of as wireframe mesh covered with texture information, or a “texel image.” Examples of sensors developed to achieve this result include tripod-based,³ mobile platform-based,⁴ and handheld configurations.⁵

There has also been several papers on Time-of-Flight (TOF) sensor calibration. These include calibration of a sensor similar to ours,⁶ and calibration issues both for geometric and intensity errors.⁷⁻⁹ Our method differs in that in our configuration, the sensors are co-boresighted and the visible image and ladar depth imaged are fused at the pixel level without parallax.

This paper describes the calibration process for texel cameras that are constructed from a digital image sensor and a flash ladar array. Although a hand-held texel camera employing a TOF ladar sensor is described, the methods can be adapted to any combination of camera sensor and flash ladar. The remainder of the paper proceeds as follows. Section 2 describes the combined sensor used in this work, while Section 3 describes the calibration for the geometric distortion caused by the lenses on the individual sensors. (Much of the method described on these sections is found in Boldt.⁵) Section 4 presents the methods used to reduce measurement errors in the TOF sensor caused by reflected intensity differences. Finally, calibration results are given in Section 5, and Section 6 concludes the paper.

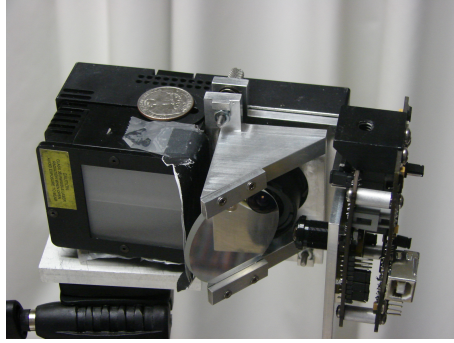


Figure 1. Texel camera constructed from a TOF depth sensor and a imaging sensor.

2. CAMERA CONFIGURATION

The camera configuration used to develop these calibration methods consists of a Micron 1280 x 1024 CMOS color imaging sensor integrated with a Canesta 64 x 64 TOF ladar sensor. (Details of the system are given in Boldt.⁵) The cameras are co-boresighted by means of a cold mirror which allows the infrared light from the ladar to pass through to the depth sensor while the visible light is reflected into the aperture of the imaging sensor. The cold mirror and the cameras are placed in a mechanical mount so that the different sensors have principle axes that are perpendicular, with the mirror at a 45° angle between them. The field-of-view of the combined sensors is 55°. A picture of the texel camera is given in Figure 1.

Before calibration can proceed, the sensors must be co-boresighted so that their fields of view overlap, and the effect of parallax is eliminated. Parallax will result if the sensors do not share a common center of perspective (COP). The center of perspective is the point about which a sensor can be rotated without a relative shift in the position of objects located at different distances in the scene.¹⁰ The position of the mirror is adjusted until the principal axes of the two sensors exit the mirror coaxially and the COP of the ladar sensor is at the same spot as the virtual COP of the imaging sensor. The process is done in four steps:

1. **Match the fields of view.** The tilt on the cold mirror is adjusted so that all points in the visual image are in focus, and the center of the image sensor field of view matches the center of the field of view of the ladar. Note that both a depth image and a “brightness” image are acquired by the ladar. It is typically easier to use the brightness image than the depth image for COP alignment.
2. **Find the COP of both the image sensor and the ladar.** This is done using a panoramic mount.¹¹ The texel camera is mounted to find the COP of the ladar. Both of the cameras can then be checked simultaneously, since the virtual COP of the image sensor is designed to be at the location of the COP of the ladar. When the ladar COP is found in either the vertical or horizontal axis, rotation about that axis will cause no shift in objects in the foreground relative to the background in the brightness image. If the virtual COP of the image sensor is located at the same point, there will be no shift in objects in the visual image during rotation.
3. **Adjust the mirror or the image sensor.** The mirror or image sensor must be translated without rotation so that the image sensor virtual COP is at the location of the ladar COP.
4. **Repeat.** Iterate steps 1–3 if necessary.

At this point, the mechanical adjustments can be fixed, and the camera is ready to be calibrated. The calibration of the texel camera includes two major steps. These consist of a geometric calibration, which is applied to both of the sensors, followed by a calibration of the ladar for brightness errors.

3. GEOMETRIC CALIBRATION

The geometric calibration is required because both of the sensors contain a lens to focus light onto their focal planes. One approach is to calibrate each of the sensors individually and find the mapping between them, however, we choose to calibrate the lidar, and then find the mapping between the image sensor and the lidar. In this way, points on the visual image can be mapped directly to corresponding points on the lidar brightness image. Since the brightness image has been calibrated, the visual image will become calibrated through the mapping.

Since the TOF lidar provides depth information, it does not create a true point cloud. The calibration must therefore be based on knowing the depth of the image from the lidar. If a flash lidar measuring range is used, the methods reported here can be easily modified to accommodate range measurements. It is assumed that the depth is measured along the -z-axis of the camera coordinate system.

3.1 Lidar Sensor Calibration

Let a point in space be defined as

$$\mathbf{P}_r = [X_r, Y_r, Z_r]. \quad (1)$$

Given an ideal pinhole camera, this point is projected onto a 2D sensor array at a location given by

$$\mathbf{p}_n = \begin{bmatrix} \frac{X_r}{Z_r} \\ \frac{Y_r}{Z_r} \end{bmatrix} = \begin{bmatrix} x_n \\ y_n \end{bmatrix}, \quad (2)$$

where \mathbf{p}_n is the normalized coordinate of the point. The actual location of the point is changed by non-ideal characteristics of the camera, including lens distortion.

Using a model of camera distortion proposed by Heikkilä and Silvén,¹² the distorted point coordinate is given by

$$\mathbf{p}_d = \begin{bmatrix} x_d \\ y_d \end{bmatrix} = d_r \begin{bmatrix} x_n \\ y_n \end{bmatrix} + \mathbf{d}_t, \quad (3)$$

where d_r is caused by radial distortion, and \mathbf{d}_t is caused by tangential distortion. These are defined by

$$d_r = 1 + k_1 r^2 + k_2 r^4 + k_5 r^6, \quad (4)$$

$$\mathbf{d}_t = \begin{bmatrix} 2k_3 x_n y_n + k_4 (r^2 + 2x_n^2) \\ k_3 (r^2 + 2y_n^2) + 2k_4 x_n y_n \end{bmatrix}, \quad (5)$$

and

$$r^2 = x_n^2 + y_n^2. \quad (6)$$

The position of the point in pixel coordinates is determined by the camera calibration parameters given in projective coordinates by

$$\begin{bmatrix} \mathbf{p}_p \\ 1 \end{bmatrix} = \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix}, \quad (7)$$

where \mathbf{p}_p is the pixel position (in pixels), and \mathbf{K} is the camera matrix containing the intrinsic parameters, where \mathbf{K} defined as:

$$\mathbf{K} = \begin{bmatrix} f_1 & 0 & cc_1 \\ 0 & f_2 & cc_2 \\ 0 & 0 & 1 \end{bmatrix}. \quad (8)$$

The parameters f_1 and f_2 are, respectively, the horizontal and vertical focal distances, and cc_1 and cc_2 are the horizontal and vertical position of the principal point with reference to the origin of the pixel matrix (such that (0,0) is located at the upper left corner).¹³

The camera parameters given in (4)–(8) must be found to calibrate the lidar. Methods such as those proposed by Zhang,¹⁴ Heikkilä and Silvén,¹⁵ or Tsai¹⁶ can be used to do this.

3.2 Calibrated 3D Points

The parameters found using the methods in Section 3.1 are used to find the true location of the 3D points measured by the lidar. The points are corrected knowing the measured depth and the distortion caused by the lidar sensor lens. Given the pixel coordinates \mathbf{p}_p and the measured depth Z_r , it is possible to compute \mathbf{P}_r . This is done as follows.

First, the camera calibration matrix is inverted and the distorted point locations \mathbf{p}_d are obtained. From these, the normalized points, \mathbf{p}_n , are found by inverting the nonlinear mapping given by (3)–(6), which can be reduced to

$$\begin{aligned} x_d(x_n, y_n) &= k_5 x_n^7 + 3 k_5 x_n^6 y_n^2 + k_2 x_n^5 + 3 k_5 x_n^3 y_n^4 + 2 k_2 x_n^3 y_n^2 + k_1 x_n^3 \\ &\quad + 3 k_4 x_n^2 + k_5 x_n y_n^6 + k_2 x_n y_n^4 + k_1 x_n y_n^2 + 2 k_3 x_n y_n + x_n + k_4 y_n^2 \\ y_d(x_n, y_n) &= k_5 x_n^6 y_n + 3 k_5 x_n^4 y_n^3 + k_2 x_n^4 y_n + 3 k_5 x_n^2 y_n^5 + 2 k_2 x_n^2 y_n^3 + k_1 x_n^2 y_n \\ &\quad + k_3 x_n^2 + 2 k_4 x_n y_n + k_5 y_n^7 + k_2 y_n^5 + k_1 y_n^3 + 3 k_3 y_n^2 + y_n. \end{aligned} \quad (9)$$

There is no analytical solution for (x_n, y_n) , though algorithms for approximating the solution have been proposed.^{15,17} Following the approach of Melen,¹⁷ Algorithm 1 can be used, where $\delta(\mathbf{p}_n) = d_r \mathbf{p}_n + \mathbf{d}_t - \mathbf{p}_n$, and d_r and \mathbf{d}_t are given in (4) and (5), respectively. Note that the solution for \mathbf{p}_n is fixed for each pixel in the lidar array, and can be pre-computed and used in a table lookup.

Algorithm 1 Iterative approach to transforming the image coordinates from distorted to normalized space.

Input:

Distorted Normalized Coordinates,
 $\mathbf{p}_d = (x_d, y_d)$,
 Radial and Tangential Distortion Coefficients,
 $k_1 - k_5$,
 Number of Iterations to Perform,
 n

Output:

Approximate Undistorted Normalized Coordinates,
 $\tilde{\mathbf{p}}_n = (\tilde{x}_n, \tilde{y}_n)$

Begin

For Each Pixel

$\tilde{\mathbf{p}}_n = \mathbf{p}_d$
Loop For n Iterations
 $\tilde{\mathbf{p}}_n = \mathbf{p}_d - \delta(\tilde{\mathbf{p}}_n)$
Next Iteration

Next Pixel

End

Finally, the corrected 3D point is given by

$$\mathbf{P}_r = [X_r, Y_r, Z_r] = [x_n Z_r, y_n Z_r, Z_r]. \quad (10)$$

3.3 Lidar-to-Image Transformation

In general, the image sensor will have its own distortion caused by optical system imperfections. For the short focal length lens used in our camera, the barrel distortion was noticeable, as shown in Figure 2 (a). This distortion must be corrected to obtain a calibrated image.

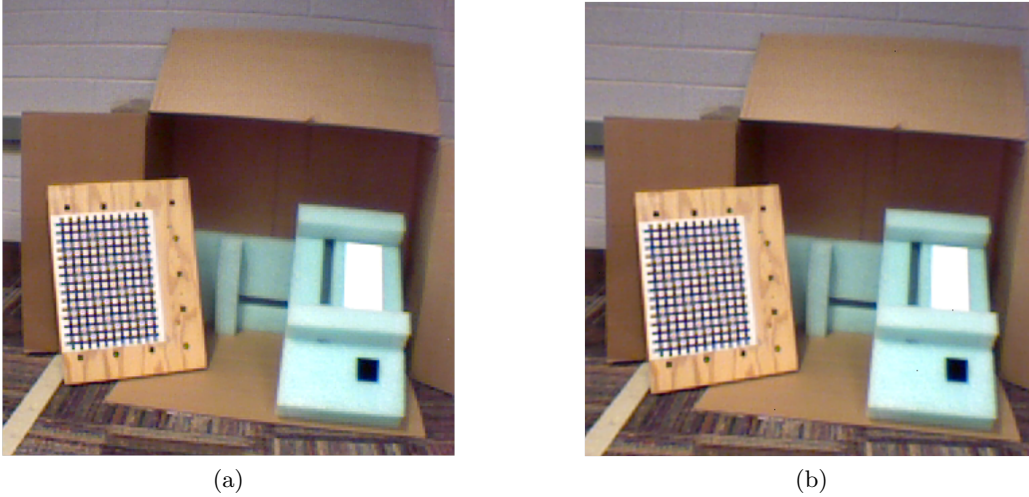


Figure 2. Examples of visual image (a) before and (b) after image-to-ladar transformation.

The transformation is chosen to be a 2D, third-order polynomial mapping¹⁸ given by

$$\begin{aligned} x &= a_x + b_x \tilde{x}_p + c_x \tilde{x}_p^2 + d_x \tilde{x}_p^3 + e_x \tilde{y}_p + f_x \tilde{x}_p \tilde{y}_p + i_x \tilde{y}_p^2 + j_x \tilde{x}_p \tilde{y}_p^2, \\ y &= a_y + b_y \tilde{x}_p + c_y \tilde{x}_p^2 + e_y \tilde{y}_p + f_y \tilde{x}_p \tilde{y}_p + g_y \tilde{y}_p \tilde{x}_p^2 + i_y \tilde{y}_p^2 + m_y \tilde{y}_p^3, \end{aligned} \quad (11)$$

where (x, y) are pixel coordinates in the visual sensor system, and $(\tilde{x}_p, \tilde{y}_p)$ are pixel coordinates in the calibrated brightness image given by

$$\begin{bmatrix} \tilde{x}_p \\ \tilde{y}_p \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \tilde{x}_n \\ \tilde{y}_n \\ 1 \end{bmatrix}. \quad (12)$$

Equation (11) includes terms for translation, rotation, scaling, tilt, and radial nonlinearities. The parameters in (11) are estimated by finding corresponding points in visible images and brightness images, and finding a least-squares fit to the points.

4. DEPTH CALIBRATION

An effort was made to provide more accurate calibration of the depth measurements from the ladar camera. It was observed that the measurement of depth on highly intensity returns and low intensity returns is not the same for surfaces positioned the same distance from the camera. This phenomenon, commonly referred to as “walk error,” should be calibrated from the camera, if possible. The most effective method to accomplish this is to create a calibration surface for each pixel in the depth image, or with groups of pixels in the image, which is a function of the returned intensity of pixel and the measured depth.^{7,9} Since the goal of this effort is to create texel images in real-time, a simpler approach was attempted.

To test a simple calibration curve, a depth image of a flat surface at about a depth of 80 cm was acquired, as shown in Figure 3. A test target consisting of a white piece of paper with a black square was mounted to the wall to provide different reflectivity regions in the image. From the images it is apparent that there is walk error present, but the error seems to be small in magnitude, except where the dark areas are located. It is also obvious that the depth camera suffers from significant vignetting.

A calibration curve was computed by finding the 3rd-order least-squares polynomial fit to the error from the measured depth as a function of the intensity of each point. The resulting curve is given in Figure 4. From the figure it is clear that the data from this depth camera is very noisy, but causes a decrease in the depth (in the -z direction) for dark pixels and increases the depth for bright pixels.

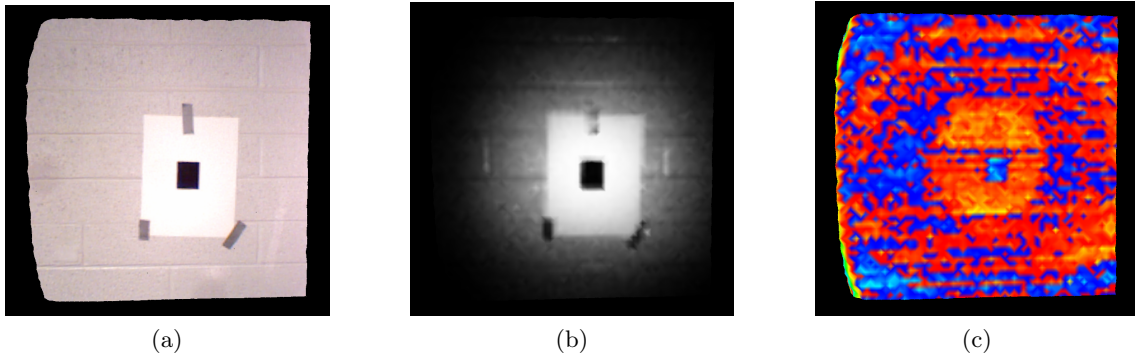


Figure 3. Images of a flat wall taken from the texelcamera. (a) Color image. (b) Intensity image. (c) Depth image, where yellow corresponds to closer areas and blue corresponds to areas farther from the camera.

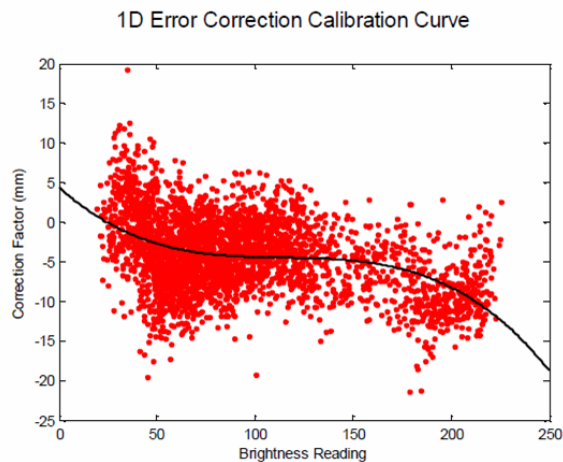


Figure 4. Calibration curve computed as a function of depth pixel intensity measurements.

5. CALIBRATION RESULTS

The texel camera shown in Figure 1 was calibrated using the methods in the previous sections. The Camera Calibration Toolbox¹⁹ was used to find the intrinsic parameters and radial and tangential distortion parameters ((4)–(8)) for the depth camera. Once these were found, Algorithm 1 was used to find the values for (x_n, y_n) . The values for these points for each of the detectors in the array are given in Figure 5. Note that the array is not strictly rectangular due to the distortion in the lens.

In order to find the lidar-to-image mapping, a checkerboard image was observed by both cameras, and corresponding points at the corners of the checkers were recorded. These point pairs were then used to find the warping parameters given in (11). Since the corrected lidar pixel points $(\tilde{x}_p, \tilde{y}_p)$ are in undistorted space, the mapping of the color image pixels are also corrected. An example of the resulting image is given in Figure 2 (b).

At the conclusion of the geometric calibration process, the color image is registered to the depth image as shown in Figure 6. The color image is registered very well with 3D points in the test scene. Corners are well-defined, and edges in the images match very well. Three-dimensional distance measurements of objects in the scene indicate errors of only about 4%.

Unfortunately, the depth calibration using the curve shown in Figure 4 did not work effectively. The main problem seems to be that for this depth camera, the error in the dark areas is difficult to characterize, and this problem is made worse by the severe vignetting observed in the depth camera.

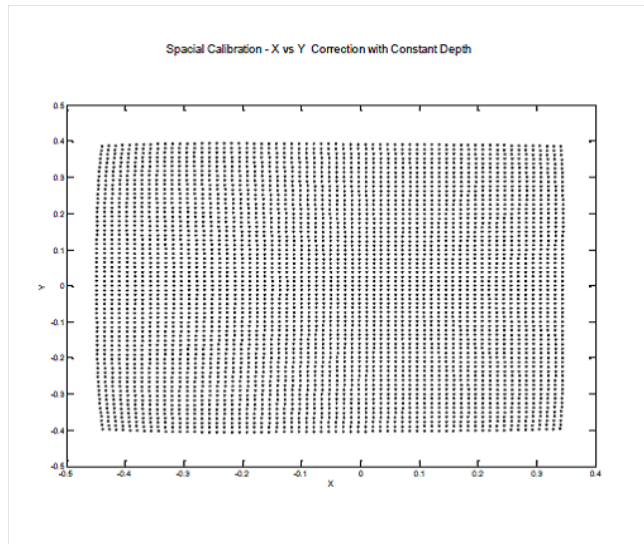
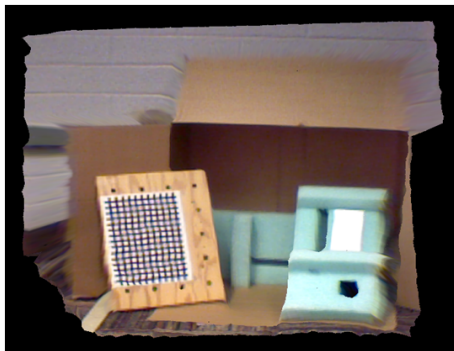
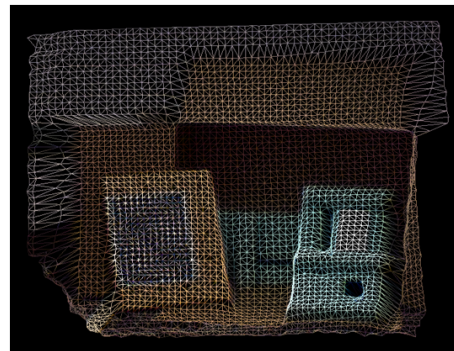


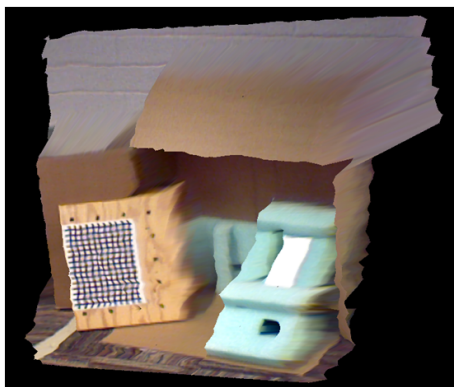
Figure 5. Corrected ladar point locations, (x_n, y_n) , found from the calibration parameters.



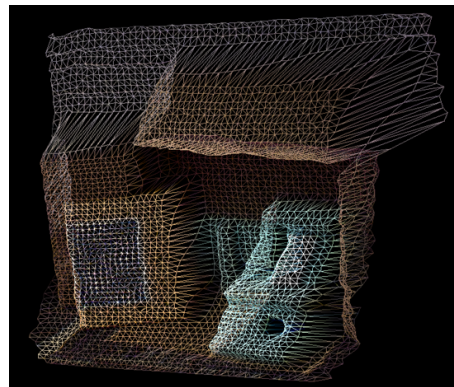
(a)



(b)



(c)



(d)

Figure 6. Examples of 3D images ((a) and (c)), and their corresponding wireframe surfaces ((b) and (d)).

6. CONCLUSION AND DISCUSSION

The methods reported in this paper were very effective in calibrating the geometry errors in the fused texel image caused by lens distortion in the individual sensors. Edges and corners match very well, and the resulting textured surface can be used to accurately measure the dimensions of objects in the scene. Since all of the calibration

parameters are computed constants, they can be used during real-time acquisition in a camera driver so that texel images are produced at near-video frame rates.

The poor performance of the depth image calibration process was unexpected. The errors seem to be a complex function of both the intensity of the lidar returns and the distance from the camera. As the TOF lidar sensor used in this work is a early-generation TOF sensor, it was determined that additional tests of the method should be done using a newer-generation TOF sensor, to confirm the value of the method for general use.

7. ACKNOWLEDGEMENTS

The author would like to acknowledge the help of Brittin Bennett and Neeraj Badamikar in the preparation of this paper.

REFERENCES

- [1] Hahne, U. and Alexa, M., “Combining Time-Of-Flight depth and stereo images without accurate extrinsic calibration,” *Int. J. Intelligent Systems Technologies and Applications* **5**(3/4), 325–333 (2008).
- [2] Zhang, J., Wang, L.-H., Li, D.-X., and Zhang, M., “High quality depth maps from stereo matching and ToF camera,” in [*Int. Conf. of Soft Computing and Pattern Recognition (SoCPaR)*], 68–72 (Oct. 2011).
- [3] Pack, R. T., Israelsen, P., and Sealy, K., “A co-boresighted synchronized lidar/EO imager for creating 3D images of dynamic scenes,” in [*Laser Radar Technology and Applications X*], Kamerman, G. W., ed., **5791**, 42–50, SPIE (May 2005).
- [4] Pack, R. T., Swasey, J. A., Fullmer, R. R., Budge, S. E., Israelsen, P. D., Petersen, B., and Cook, T. D., “Eyesafe LADAR test-bed with coaxial color imager,” in [*Laser Radar Technology and Applications XIV*], Turner, M. D. and Kamerman, G. W., eds., **7323**, 732303, SPIE, Orlando, FL, USA (Apr. 2009).
- [5] Boldt, B. M., Budge, S. E., Pack, R. T., and Israelsen, P. D., “A handheld texel camera for acquiring near-instantaneous 3D images,” in [*Proc. Asilomar Conf. Signals, Systems, and Computers*], (Nov. 2007).
- [6] Lindner, M., Kolb, A., and Hartmann, K., “Data-fusion of PMD-based distance-information and high-resolution RGB-images,” in [*Int. Symp. on Signals, Circuits and Systems (ISSCS)*], **1**, 1–4 (july 2007).
- [7] Lindner, M. and Kolb, A., “Calibration of the intensity-related distance error of the PMD TOF-camera,” *Intelligent Robots and Computer Vision XXV: Algorithms, Techniques, and Active Vision* **6764**(1), 67640W, SPIE (2007).
- [8] Kahlmann, T., Remondino, F., and Ingensand, H., “Calibration for increased accuracy of the range imaging camera Swissranger,” in [*Image Engineering and Vision Metrology (IEVM)*], 136–141 (2006).
- [9] Lindner, M. and Kolb, A., “Lateral and depth calibration of PMD-distance sensors,” in [*Int. Symp. on Visual Computing (ISVC)*], Bebis, G. et al., eds., 524–533, Springer-Verlag (2006).
- [10] Altenhofen, R. E. and Hedden, R. T., “Transformation and rectification,” in [*Manual of Photogrammetry*], Thompson, M. M., ed., *Advances in Electronics and Electron Physics* **2**, 1–59, American Society of Photogrammetry (1966).
- [11] Milburn, K., [*Digital Photography: Expert Techniques*], O’Reilly & Associates, Inc. (2003).
- [12] Heikkilä, J. and Silvén, O., “Calibration procedure for short focal length off-the-shelf CCD cameras,” in [*Proc. of the 13th Int. Conf. on Pattern Recognition*], **1**, 166–170 (1996).
- [13] Ma, Y., Soatto, S., Kosecka, J., and Sastry, S. S., [*An Invitation to 3-D Vision: From Images to Geometric Models*], SpringerVerlag, New York City (2003).
- [14] Zhang, Z., “Flexible camera calibration by viewing a plane from unknown orientations,” in [*The 7th IEEE Int. Conf. Computer Vision*], **1**, 666–673, IEEE (Sept. 1999).
- [15] Heikkilä, J. and Silvén, O., “A four-step calibration procedure with implicit image correction,” in [*Proc. IEEE Conf. Computer Vision and Pattern Recognition*], (1997).
- [16] Tsai, R. Y., “A versatile camera callibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses,” *IEEE Trans. Robot. Autom.* **3**, 323–344 (Aug. 1987).
- [17] Melen, T., *Geometrical Modeling and Calibration of Video Cameras for Underwater Navigation*, dr. Ing. thesis, Norges Tekniske Høgskole, Institutt for Teknisk Kybernetikk (1994).
- [18] Mayampurath, A. M., *Real-Time Cell Image Analysis*, Master’s thesis, Utah State University, Logan (2005).
- [19] Bouguet, J.-Y., “Camera calibration toolbox.” Web URL http://www.vision.caltech.edu/bouguetj/calib_doc/.