Optical Engineering

SPIEDigitalLibrary.org/oe

Calibration method for texel images created from fused flash lidar and digital camera images

Scott E. Budge
Neeraj S. Badamikar

SPIE

Calibration method for texel images created from fused flash lidar and digital camera images

Scott E. Budge
Neeraj S. Badamikar
Utah State University, Center for Advanced Imaging Ladar, Dept. of Electrical and Computer Engineering, 4120 Old Main Hill, Logan, Utah 84321-4120
E-mail: scott.budge@ece.usu.edu

Abstract. The fusion of imaging lidar information and digital imagery results in 2.5-dimensional surfaces covered with texture information, called texel images. These data sets, when taken from different viewpoints, can be combined to create three-dimensional (3-D) images of buildings, vehicles, or other objects. This paper presents a procedure for calibration, error correction, and fusing of flash lidar and digital camera information from a single sensor configuration to create accurate texel images. A brief description of a prototype sensor is given, along with a calibration technique used with the sensor, which is applicable to other flash lidar/digital image sensor systems. The method combines systematic error correction of the flash lidar data, correction for lens distortion of the digital camera and flash lidar images, and fusion of the lidar 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. Results with the prototype sensor show very good match between 3-D points and the digital image (<2.8 image pixels), with a 3-D object measurement error of <0.5%, compared to a noncalibrated error of ~3%. [DOI: 10.1117/1.OE.52.10.103101]

Subject terms: lidar; ladar; calibration; sensor fusion; texel camera.

Paper 130905P received Jun. 18, 2013; revised manuscript received Aug. 29, 2013; accepted for publication Sep. 6, 2013; published online Oct. 3, 2013.

1 Introduction

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

There has been much interest in combining information obtained from multiple sensors at a single location in order to create 2.5-dimensional (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 accurately in areas of the images with low detail or low contrast. One solution is to add a range sensor (such as a lidar) 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 digital [electro-optic (EO)] camera and lidar to create a 2.5-D image. If the lidar 3-D information is fused to the EO 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, called texel cameras, include tripod-based,[3] mobile platform-based,[4] and handheld configurations.[5]

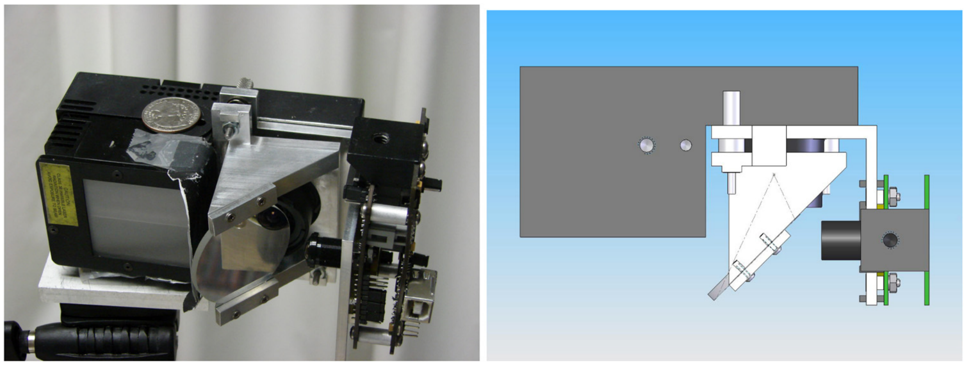
0091-3286/2013/$25.00 © 2013 SPIE

There have also been several papers on time-of-flight (TOF) sensor calibration. These include calibration of a sensor similar to ours,[6] fundamental calibration issues[7,8] and calibration for geometric, range, and brightness errors.[9–13] Reynolds et al.[14] and Frank et al.[15] investigated the problem of confidence measures to eliminate bad pixels in a TOF range image. Our method differs from the existing methods because the sensors are coboresighted and the EO image and lidar range image are fused at the pixel level without parallax.

Since two sensors are combined in a single camera, it is necessary to calibrate both sensors and determine the mapping between them to obtain accurate texel images. The calibration process must take into account the distortions caused by the lenses (if any) in both the lidar sensor and the EO camera, the mapping between the two, and the systematic errors in the lidar measurements.

This paper describes the calibration process for texel cameras that are constructed from an EO image sensor and a flash lidar array. Although a hand-held texel camera employing a TOF lidar sensor is described, the methods can be adapted to any combination of EO sensor and flash lidar. However, the wiggling and walk error correction described in Sec. 4.1 may not apply to pulse TOF flash lidars.

The remainder of the paper proceeds as follows. Section 2 describes the combined sensor used in this work, while Sec. 3 describes the calibration for the geometric distortion caused by the lenses on the individual sensors. (Much of the method described in Sec. 2 is found in Ref. 16.) Section 4 presents the methods used to reduce range measurement errors in the TOF sensor. Finally, calibration results are given in Sec. 5, and Sec. 6 concludes the paper.



(a) (b)

Fig. 1 Texel camera constructed from a time-of-flight depth sensor and an imaging sensor. (a) Photograph of prototype. (b) Solid model of design.

2 Camera Configuration

The camera configuration used to develop these calibration methods consists of a Micron 1280×1024 CMOS color EO sensor integrated with a Canesta 64×64 TOF lidar sensor. (Details of the system are given in Ref. 5.) The cameras are co-boresighted by means of a cold mirror, which allows the infrared light from the lidar to pass through to the range sensor while the visible light is reflected into the aperture of the EO 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-deg angle between them. The field-of-view of the combined sensors is 44 deg. A picture of the texel camera is given in Fig. 1, along with a solid model of the design.

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 COP 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.[17] The position of the mirror is adjusted until the principal axes of the two sensors exit the mirror coaxially and the COP of the lidar sensor is at the same spot as the virtual COP of the EO 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 EO image are in focus and the center of the EO sensor field of view matches the center of the field of view of the lidar. Note that both a range image and a brightness image are acquired by the lidar. It is typically easier to use the brightness image than the range image for COP alignment.
2. Find the COP of both the image sensor and the lidar: This is done using a panoramic mount.[18] The texel camera is mounted to find the COP of the lidar sensor. Both the lidar sensor and the EO sensor comprising the texel camera can then be checked simultaneously, since the virtual COP of the EO sensor is designed to be at the location of the COP of the lidar. When the lidar 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 EO sensor is located at the same point, there will be no shift in objects in the EO image during rotation.
3. Adjust the mirror or the image sensor. The mirror or EO sensor must be translated and/or rotated so that the EO sensor virtual COP is at the location of the lidar COP.
4. Repeat: Iterate steps 1 to 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 lidar for range errors.

3 Geometric Calibration

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 EO sensor and the lidar. In this way, points on the EO image can be mapped directly to corresponding points on the lidar brightness image. Since the brightness image has been calibrated, the EO image will become calibrated through the mapping.

In the following, it is assumed that the depth, Z_r, is measured along the principle axis (z-axis) of the texel camera coordinate system.

3.1 *Ladar Sensor Calibration Model*

Following the development given by Boldt et al.,[5] let a point in space be defined as $\boldsymbol{P}_r = [X_r, Y_r, Z_r]^T$. Given an ideal pinhole camera, this point is projected onto a two-dimensional (2-D) plane at a location given by

$$\boldsymbol{p}_n = \begin{bmatrix} \frac{X_r}{Z_r} \\ \frac{Y_r}{Z_r} \end{bmatrix} = \begin{bmatrix} x_n \\ y_n \end{bmatrix}, \tag{1}$$

where $\boldsymbol{p}_n$ is the normalized coordinate of the point. The actual location of the point is changed by nonideal characteristics of the camera, including lens distortion.

Using a model of camera distortion proposed by Heikkilä and Silvén,[19] the distorted point coordinate is given by

$$\boldsymbol{p}_d = \begin{bmatrix} x_d \\ y_d \end{bmatrix} = d_r \begin{bmatrix} x_n \\ y_n \end{bmatrix} + \boldsymbol{d}_t, \tag{2}$$

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

$$d_r = 1 + k_1 r^2 + k_2 r^4 + k_5 r^6, \tag{3}$$

$$\boldsymbol{d}_t = \begin{bmatrix} 2k_3 x_n y_n + k_4(r^2 + 2x_n^2) \\ k_3(r^2 + 2y_n^2) + 2k_4 x_n y_n \end{bmatrix}, \tag{4}$$

$$r^2 = x_n^2 + y_n^2. \tag{5}$$

The position of the point in the sensor array in pixel coordinates is determined using the camera calibration parameters and the distorted points in homogeneous form using

$$\boldsymbol{p}_p = \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = \boldsymbol{K} \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix}, \tag{6}$$

where $\boldsymbol{p}_p$ is the pixel position and $\boldsymbol{K}$ is the camera matrix containing the intrinsic parameters.

$$\boldsymbol{K} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}. \tag{7}$$

The parameters f_x and f_y are, respectively, the horizontal and vertical focal distances, s is the detector skew, and c_x and c_y are the horizontal and vertical positions of the principal point with reference to the origin of the pixel matrix [such that (0,0) is located at the upper left corner].[20]

The camera parameters given in Eqs. (3), (4), and (7) must be found to calibrate the lidar. Methods such as those proposed by Zhang,[21] Heikkilä and Silvén,[22] or Tsai[23] can be used to do this.

3.2 *Calibrated 3-D Points*

The parameters found using the method in Sec. 3.1 are used to find the true location of the 3-D points measured by the lidar. The points are corrected knowing the measured range and the distortion caused by the lidar sensor lens. Given the pixel coordinates $\boldsymbol{p}_p$ and the measured range λ_r, it is possible to compute $\boldsymbol{P}_r$. This is done as follows.

First, the camera matrix $\boldsymbol{K}$ is inverted and the distorted point locations $\boldsymbol{p}_d$ are obtained. From these, the normalized points, $\boldsymbol{p}_n$, are found by inverting the nonlinear mapping given by Eqs. (2) to (5), which can be written as

$$\begin{aligned} &x_d(x_n, y_n) \\ &= x_n + k_1 x_n r^2 + k_2 x_n r^4 + k_5 x_n r^6 + 2k_3 x_n y_n + k_4(r^2 + 2x_n^2) \\ &y_d(x_n, y_n) \\ &= y_n + k_1 y_n r^2 + k_2 y_n r^4 + k_5 y_n r^6 + k_3(r^2 + 2y_n^2) + 2k_4 x_n y_n. \end{aligned} \tag{8}$$

There is no analytical solution for (x_n, y_n), though algorithms for approximating the solution have been proposed.[22,24]

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

Input:

Distorted normalized coordinates,

$p_d = (x_d, y_d)$,

Radial and tangential distortion coefficients,

$k_1 - k_5$,

Number of iterations to perform,

m

Output:

Approximate undistorted normalized coordinates,

$\tilde{\boldsymbol{p}}_n = (\tilde{x}_n, \tilde{y}_n)$

Begin

For each pixel

$\tilde{\boldsymbol{p}}_n = \boldsymbol{p}_d$

Loop For m **Iterations**

$\tilde{\boldsymbol{p}}_n = \boldsymbol{p}_d - \delta(\tilde{\boldsymbol{p}}_n)$

Next Iteration

Next Pixel

End

Following the approach of Melen,[24] Algorithm 1 can be used, where $\delta(\boldsymbol{p}_n) = d_r \boldsymbol{p}_n + \boldsymbol{d}_t - \boldsymbol{p}_n$, and d_r and $\boldsymbol{d}_t$ are given in Eqs. (3) and (4), respectively. Note that the solution for $\boldsymbol{p}_n$ is fixed for each pixel in the lidar array and can be precomputed and used in a look-up table (LUT).

Once the calibration parameter estimates $\tilde{\boldsymbol{p}}_n$ are known, it is straightforward to compute the 3-D point from the measured range λ_r. The 3-D point can be computed directly using[25]

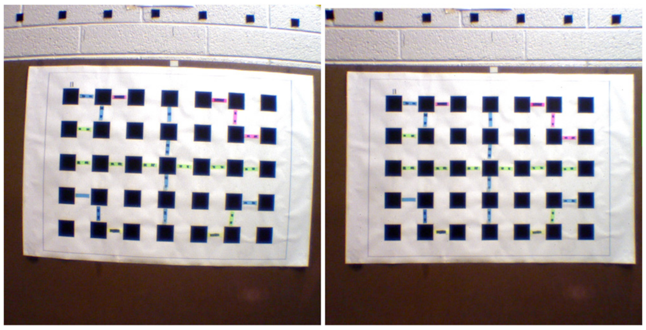
$$\tilde{\boldsymbol{P}}_r = \frac{\lambda_r}{(\tilde{x}_n^2 + \tilde{y}_n^2 + 1)^{\frac{1}{2}}} \begin{bmatrix} \tilde{x}_n \\ \tilde{y}_n \\ 1 \end{bmatrix} = \lambda_r \begin{bmatrix} \tilde{x}_n z_c \\ \tilde{y}_n z_c \\ z_c \end{bmatrix} = \lambda_r \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}, \tag{9}$$

where

$$z_c = (\tilde{x}_n^2 + \tilde{y}_n^2 + 1)^{-\frac{1}{2}}. \tag{10}$$

3.3 *Lidar-to-Image Transformation*

In general, the EO 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 Fig. 2(a). This distortion must be corrected to obtain a calibrated image.



(a) (b)

Fig. 2 Examples of electro-optic image (a) before and (b) after image-to-lidar transformation.

The transformation is based on an approach using the camera model, given by Eq. (8), applied to the pixels in the normalized pixel space $\boldsymbol{p}_n$ and corresponding pixel coordinates in the EO sensor system, given by (u, v). Using lidar pixels in the normalized space, we remove the effects of the lidar sensor lens (a pinhole model), and the EO pixels generated from this transform will naturally be square. If we apply an unknown camera matrix $\boldsymbol{K}^I$ to the camera model for the lidar sensor, the mapping becomes

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \boldsymbol{K}^I \begin{bmatrix} \tilde{x}_d(\tilde{x}_n, \tilde{y}_n) \\ \tilde{y}_d(\tilde{x}_n, \tilde{y}_n) \\ 1 \end{bmatrix} = \begin{bmatrix} f_x^I \tilde{x}_d(\tilde{x}_n, \tilde{y}_n) + s^I \tilde{y}_d(\tilde{x}_n, \tilde{y}_n) + c_x^I \\ f_y^I \tilde{y}_d(\tilde{x}_n, \tilde{y}_n) + c_y^I \\ 1 \end{bmatrix}. \tag{11}$$

After combining constants, the resulting mapping is given by

$$\begin{aligned} u &= g_1 \tilde{x}_n + g_2 \tilde{x}_n r^2 + g_3 \tilde{x}_n r^4 + g_4 \tilde{x}_n r^6 + g_5 \tilde{x}_n \tilde{y}_n \\ &\quad + g_6 (r^2 + 2\tilde{x}_n^2) + g_7 \tilde{y}_n + g_8 \tilde{y}_n r^2 + g_9 \tilde{y}_n r^4 \\ &\quad + g_{10} \tilde{y}_n r^6 + g_{11} (r^2 + 2\tilde{y}_n^2) + g_{12} v \\ &= h_1 \tilde{y}_n + h_2 \tilde{y}_n r^2 + h_3 \tilde{y}_n r^4 + h_4 \tilde{y}_n r^6 + h_5 (r^2 + 2\tilde{y}_n^2) \\ &\quad + h_6 \tilde{x}_n \tilde{y}_n + h_7. \end{aligned} \tag{12}$$

There are two potential problems with this mapping. First, the high order of this polynomial required by the r^6 term in Eq. (3) can lead to an ill-conditioned solution, and second, the constraint that the radial distortion is strictly circular may be unreasonable since imperfect alignment of the cold mirror may cause any radial distortion to be slightly eccentric. Therefore, a more appropriate mapping is formulated by removing the g_4, g_{10}, and h_4 terms, decoupling terms to relax the dependence on r^2, and rewriting to get

$$\begin{aligned} u &= g_1 + g_2 \tilde{x}_n + g_3 \tilde{x}_n^2 + g_4 \tilde{x}_n^3 + g_5 \tilde{y}_n + g_6 \tilde{x}_n \tilde{y}_n + g_7 \tilde{y}_n^2 \\ &\quad + g_8 \tilde{y}_n^3 + g_9 \tilde{x}_n^2 \tilde{y}_n + g_{10} \tilde{x}_n \tilde{y}_n^2 + g_{11} \tilde{x}_n^5 + g_{12} \tilde{y}_n^5 + g_{13} \tilde{x}_n \tilde{y}_n^4 \\ &\quad + g_{14} \tilde{x}_n^4 \tilde{y}_n + g_{15} \tilde{x}_n^3 \tilde{y}_n^2 + g_{16} \tilde{x}_n^2 \tilde{y}_n^3 \\ v &= h_1 + h_2 \tilde{y}_n + h_3 \tilde{y}_n^2 + h_4 \tilde{y}_n^3 + h_5 \tilde{y}_n \tilde{x}_n + h_6 \tilde{x}_n^2 + h_7 \tilde{y}_n \tilde{x}_n^2 \\ &\quad + h_8 \tilde{y}_n^5 + h_9 \tilde{y}_n \tilde{x}_n^4 + h_{10} \tilde{y}_n^3 \tilde{x}_n^2. \end{aligned} \tag{13}$$

Finally, since the skew parameter s is negligible for nearly all focal planes, and to allow for rotation between the lidar and EO sensors, the mapping can be reduced to

$$\begin{aligned} u &= g_1 + g_2 \tilde{x}_n + g_3 \tilde{x}_n^2 + g_4 \tilde{x}_n^3 + g_5 \tilde{y}_n + g_6 \tilde{x}_n \tilde{y}_n + g_7 \tilde{y}_n^2 \\ &\quad + g_8 \tilde{x}_n \tilde{y}_n^2 + g_9 \tilde{x}_n^5 + g_{10} \tilde{x}_n \tilde{y}_n^4 + g_{11} \tilde{x}_n^3 \tilde{y}_n^2 \\ v &= h_1 + h_2 \tilde{y}_n + h_3 \tilde{y}_n^2 + h_4 \tilde{y}_n^3 + h_5 \tilde{x}_n + h_6 \tilde{y}_n \tilde{x}_n + h_7 \tilde{x}_n^2 \\ &\quad + h_8 \tilde{y}_n \tilde{x}_n^2 + h_9 \tilde{y}_n^5 + h_{10} \tilde{y}_n \tilde{x}_n^4 + h_{11} \tilde{y}_n^3 \tilde{x}_n^2. \end{aligned} \tag{14}$$

The parameters are estimated by finding N corresponding points in EO images and brightness images ($\boldsymbol{x}_i \leftrightarrow \tilde{\boldsymbol{p}}_{ni}$) and finding a maximum-likelihood fit to the points, where the mapping parameters are given by

$$\begin{aligned} (\hat{\boldsymbol{g}}, \hat{\boldsymbol{h}}) &= \underset{(g,h)}{\operatorname{argmin}} \sum_{i=0}^{N-1} d(\tilde{\boldsymbol{p}}_{ni}, \hat{\tilde{\boldsymbol{p}}}_{ni})^2 \frac{\sigma_{\mathrm{EO}}}{\sigma_{\mathrm{lidar}}} + d(\boldsymbol{x}_i, \hat{\boldsymbol{x}}_i)^2, \\ \hat{\boldsymbol{x}} &= \begin{bmatrix} \hat{u} \\ \hat{v} \end{bmatrix} = F(\hat{\tilde{\boldsymbol{p}}}_n, \boldsymbol{g}, \boldsymbol{h}), \end{aligned} \tag{15}$$

where $\hat{\tilde{\boldsymbol{p}}}_{ni}$ and $\hat{\boldsymbol{x}}_i$ are the estimated positions of the correct (noiseless) observations, $\sigma_{\mathrm{EO}}/\sigma_{\mathrm{lidar}}$ is the ratio of EO image pixel size to lidar pixel size, $d(\cdot)$ is the Euclidean distance, and $F(\tilde{\boldsymbol{p}}_n, \boldsymbol{g}, \boldsymbol{h})$ is given by Eq. (14).[26]

4 Range Calibration

It is well known that systematic error in the range measurements of TOF cameras is caused by nonideal modulation functions used for measuring range.[7] This error is commonly called wiggling error because of oscillations in the error as a function of range. In addition, it is observed that the measurement of range on high brightness returns and low brightness returns is not the same for surfaces positioned the same distance from the camera. This phenomenon, commonly referred to as walk error, should also be calibrated from the measurements, if possible. An effective method to accomplish this is to create a 2-D calibration surface, which is a polynomial function of the returned brightness of a pixel and the measured range.[11] Lindner et al. added additional terms to model the radial brightness decrease caused by vignetting.[9] Since the goal of this effort is to create texel images in real time, a simpler approach was attempted.

4.1 *Range Error Correction*

First, it was observed that in our lidar sensor, there are slight differences in the location of the depth (z-axis) origin for each pixel. These differences can be explained by hardware timing and response differences experienced by each pixel. Measurements for a number of depths from a reference point on the lidar sensor to a flat wall were made, and the maximum-likelihood corrections for each pixel were computed. (This is similar to a flat-field correction of the measured depths.) These errors were converted to range corrections for each pixel, stored in an LUT, and added to the measured range for each pixel.

Since the mechanism for wiggling and walk error is a function of the nonideal modulation used in each sensor measurement, and each pixel in the sensor sees the same

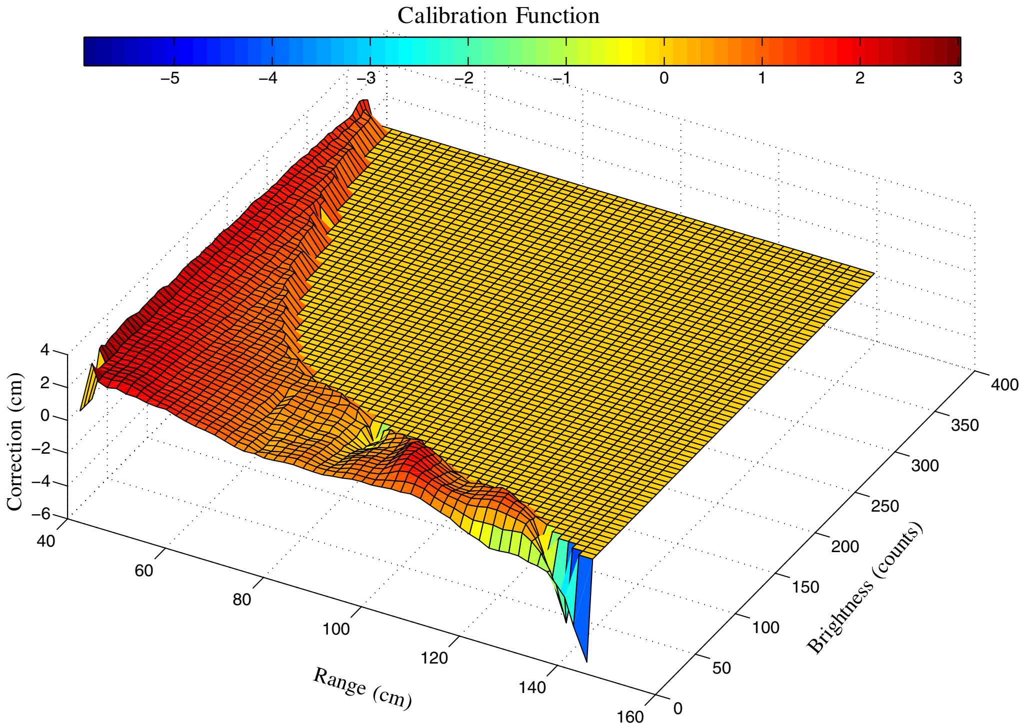


Fig. 3 Calibration function surface. The calibrated range λ_r^{cal} is corrected according to the range λ_r and brightness B for all (m, n). The flat portion of the surface is caused by points that are not possible with the Canesta lidar sensor.

modulation, a single correction function should apply to all pixels.

The range error for a dense set of range-brightness measurements (from depth and brightness images) can be used to create a calibration LUT. This is done by 2-D quantizing the calibration data and averaging the error between the measured and true ranges for all of the measurements in each quantization cell. The size of the quantization cells is based on the tradeoff between the size of the LUT and the accuracy of the error correction.

An example of an error correction function is given in Fig. 3. The range $[\lambda_r(m,n)]$ and brightness $[B(m,n)]$ are used as inputs to the LUT, and the correction to the range is found by planar interpolation of the three closest entries in the table. Once the correction is found, it is added to $\lambda_r(m,n)$ to get the calibrated range $\lambda_r^{\mathrm{cal}}(m,n)$. The wiggling error is evident in the figure as crests and troughs in the correction surface.

4.2 *COP Offset Error*

An additional source of error in the 3-D point locations is caused by the possibility that the origin for the coordinate system assumed by the lidar sensor internal processing might not be at the COP of the sensor. For a geometrically calibrated sensor, a COP with a positive Z_r value (the COP is in front of the z-coordinate origin assumed by the sensor) will cause the computed 3-D points to be farther apart than the truth, and a COP with a negative Z_r will cause the 3-D points to be closer together than the truth. The correct value of the offset (z_o) between the COP and the assumed sensor origin will be where the distance between 3-D points obtained from the sensor are equal to the distance between 3-D points measured from a physical test fixture. Since the true z_o is unknown, only the (X_r, Y_r) values are used to estimate z_o.

If we define 2-D points measured on a flat checkerboard test fixture to be $\boldsymbol{P}_r^{m,k}$, $k = 0, \ldots, K$, the fixture is located at $l = 0, \ldots, L$ different depths parallel to the lidar sensor focal plane, and the computed 2-D points from the lidar sensor are given by

$$\tilde{\boldsymbol{P}}_r^{l,k} = \begin{bmatrix} x_c^k \lambda_r^{l,k} + z_o \tilde{x}_n^k \\ y_c^k \lambda_r^{l,k} + z_o \tilde{y}_n^k \end{bmatrix}. \tag{16}$$

The estimated value of z_o is given by

$$z_o = \underset{z_o}{\operatorname{argmin}} \sum_{l=0}^{L} \sum_{j=0}^{K_l} \sum_{k=j+1}^{K_l} [d(\tilde{\boldsymbol{P}}_r^{l,j}, \tilde{\boldsymbol{P}}_r^{l,k}) - d(\boldsymbol{P}_r^{m,j}, \boldsymbol{P}_r^{m,k})]^2. \tag{17}$$

Thus, the final calibrated 3-D points are given by

$$\tilde{\boldsymbol{P}}_r(m,n) = \begin{bmatrix} X_r^{\mathrm{cal}}(m,n) \\ Y_r^{\mathrm{cal}}(m,n) \\ Z_r^{\mathrm{cal}}(m,n) \end{bmatrix} = \begin{bmatrix} x_c(m,n)\lambda_r^{\mathrm{cal}}(m,n) + z_o \tilde{x}_n(m,n) \\ y_c(m,n)\lambda_r^{\mathrm{cal}}(m,n) + z_o \tilde{y}_n(m,n) \\ z_c(m,n)\lambda_r^{\mathrm{cal}}(m,n) + z_o \end{bmatrix}. \tag{18}$$

5 Calibration Results

The texel camera shown in Fig. 1 was calibrated using the methods in the previous sections. The Camera Calibration Toolbox[27] was used to find the intrinsic parameters and radial and tangential distortion parameters [Eqs. (3) to (7)] for the

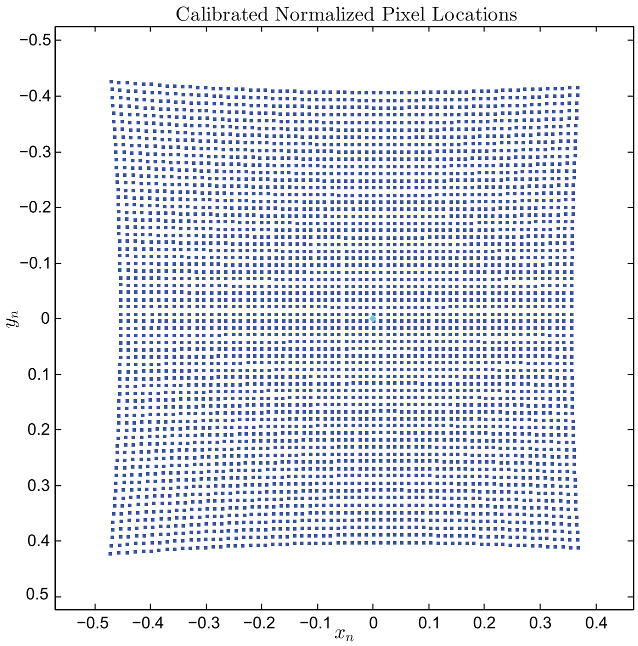


Fig. 4 Corrected lidar point locations, (x_n, y_n), found from the calibration parameters. The cyan diamond represents the principle point in the focal plane of the lidar array.

lidar 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 shown in Fig. 4. Note that the array is not strictly rectangular due to the distortion in the lens.

Results of the calibration of the lidar camera are given in Table 1.

To find the lidar-to-image mapping, an image with regularly spaced squares was observed by both lidar and EO sensors, and corresponding points at the corners of the squares were recorded. These point pairs were then used to find the warping parameters given in Eq. (14). Since the corrected normalized lidar points $(\tilde{x}_n, \tilde{y}_n)$ are in undistorted space, the mapping of the EO image pixels are also corrected. An example of the resulting image is given in Fig. 2(b).

Reprojection errors for different lidar-to-image mapping models, including those given in Eqs. (12) to (14), are listed in Table 2. As shown, the selected mapping model [Eq. (14)] provides the lowest error of the models proposed. At the conclusion of the geometric calibration process, the EO image is registered to the lidar image as shown in Fig. 5. Qualitative visual evaluation indicates that the EO image is registered very well with 3-D points in the test scene, suggesting that the 2.8-pixel reprojection error results in a good visual registration of the data sets.

Calibration data were gathered by capturing range and brightness images of a constant, high-reflectivity wall with the camera placed at depths every 2 cm within the expected working range of 30 to 140 cm. At each depth, six light source power settings were used, ranging from the minimum power needed to measure 75% of the pixels above a threshold of 14 counts, and the maximum power available from the light source before saturation. This process acquired a dense set of range-brightness pairs. Three hundred frames at each setting were averaged to reduce the effect of random noise obscuring the structured error. The resulting calibration surface is given in Fig. 3.

The value of the correction for brightness differences is illustrated in Fig. 6, where a set of eight different reflectance bars are placed on a flat wall.

Comparing Figs. 6(c) and 6(d), it is observed that the correction helped to reduce the walk error caused by the difference in reflectivity of the bars.

The accuracy of the point cloud captured from the texel camera was measured by comparing the 3-D points given by the camera and the same points measured on a fixture, as described in Sec. 4.2. The distances to the fixture were 50, 70, 90, and 110 cm. The contribution of each calibration step was compared by measuring the mean and standard deviation of the error in the measured distance between points. One thousand images were averaged to reduce random measurement error.

Table 1 Calibration parameters for the lidar camera.

Parameter	Value	3σ error bounds
Focal distance (f_x, f_y)	(80.4527, 80.3708)	$\pm(0.2318, 0.24581)$
Principal point (c_x, c_y)	(34.8945, 31.6703)	$\pm(0.3637, 0.3564)$
Skew (s)	0.00000	± 0.00000
Distortion:		
k_1	−0.19969	± 0.01315
k_2	0.05126	± 0.04457
k_3	−0.00077	± 0.00079
k_4	0.00411	± 0.00083
k_5	0	0.00000
Pixel error (σ_x, σ_y)	(0.06693, 0.06836)	

Table 2 Reprojection error for different mapping models. Errors are in electro-optic pixels.

Parameter model	Root mean square reprojection error
19-parameter [Eq. (12)]	5.6522
16-parameter [Eq. (12) without r^6 terms]	5.6593
13-parameter [Eq. (12) without r^4 and r^6 terms]	5.6654
26-parameter [Eq. (13)]	5.4895
17-parameter [Eq. (13) without r^4 terms]	5.5042
22-parameter [Eq. (14)]	2.7781
16-parameter (as in Ref. 28)	2.7991

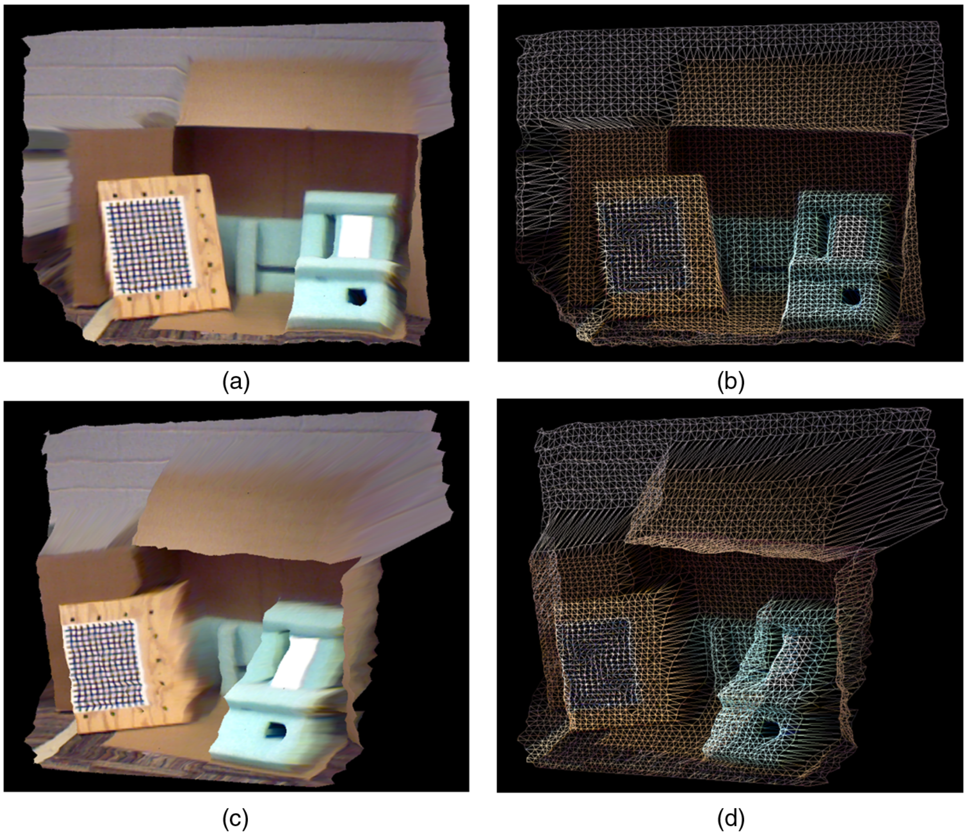


Fig. 5 Examples of texel images [(a) and (c)] and their corresponding wireframe surfaces [(b) and (d)].

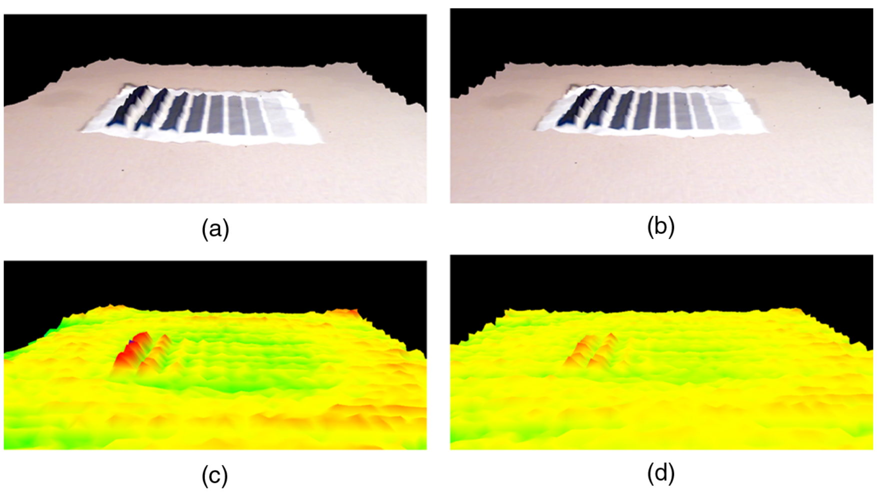


Fig. 6 Range correction on texel images (a) before and (b) after using a calibration look-up table. The corresponding three-dimensional surfaces are given in (c) and (d), colored by height.

The results are given in Table 3. The first four columns show the error without ($k_1, \ldots, k_5$ were set to zero) and with nonlinear correction. This indicates that correction for nonlinearities [Eq. (2)] reduces the average error by ~30%. The next two columns report the error with nonlinear correction and an estimated value of $z_o = 2.98$ cm (toward the focal plane). Estimation of z_0 causes a significant reduction in error to ~3% of the original average error. The final two columns report the error with all of the calibration steps, including flat-field correction and range error correction using the LUT (Sec. 4.1). In this case, the average offset found during flat-field correction was 2.0 cm with a pixel variance of 0.23 cm. This offset was included in the measured pixel data before the application of the range calibration function (Fig. 3) and estimation of z_0. After flat-field correction and range calibration, the new z_o was found to be 0.87 mm, indicating that the reference point chosen for flat-field correction was very close to the COP. The final error was only 1.26% of the original average error. From the table, it is clear that as each step of the calibration process

Table 3 Point cloud measurement error after each calibration step. All measurements are in cm.

	$z_o = 0$				$z_o = 2.98$ cm		$z_o = 0.87$ mm	
	$k_1, \ldots, k_5 = 0$		$k_1, \ldots, k_5$ from calibration		$k_1, \ldots, k_5$ from calibration		Range calibrated $k_1, \ldots, k_5$ from calibration	
Depth (Z_r)	μ	σ	μ	σ	μ	σ	μ	σ
50	2.222	1.540	1.715	0.845	0.287	0.276	0.026	0.198
70	1.398	0.567	1.067	0.406	0.013	0.276	0.143	0.270
90	1.294	0.523	0.797	0.420	−0.293	0.371	−0.113	0.276
110	1.414	0.604	0.884	0.535	−0.180	0.356	0.025	0.245
Overall	1.582	0.796	1.116	0.671	−0.043	0.385	0.020	0.260

is applied, the average error and standard deviation for all of the points decrease.

6 Conclusion

The methods reported in this paper were very effective in calibrating the errors in the fused texel image caused by lens distortion in the individual sensors. Although the intended application of this texel camera requires only a maximum range of ~1.5 m, error in the dimensions of 3-D objects are accurate to <3 mm (with low-noise measurements). For the 3-D points measured, this corresponds to <0.5% error.

The match between the fused point-cloud surface and the EO image is within an error of a few EO pixels, and the resulting textured surface can be used to accurately measure the dimensions of objects in the scene. Since the fusion of the point cloud and the EO image are performed at acquisition, calibrated texel images open up new possibilities for real-time object identification and ATR, as well as 3-D image creation. 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 calibration steps reported here include correction for sources of error that are unique to TOF lidars (i.e., wiggle error), but most of the calibration will apply to texel cameras created from any flash lidar combined with an EO sensor (including infrared and multispectral sensors).

References

1. U. Hahne and M. Alexa, "Combining time-of-flight depth and stereo images without accurate extrinsic calibration," *Int. J. Intell. Syst. Technol. Appl.* **5**(3/4), 325–333 (2008).
2. J. Zhang et al., "High quality depth maps from stereo matching and ToF camera," in *2011 Int. Conf. of Soft Computing and Pattern Recognition*, pp. 68–72, IEEE (2011).
3. R. T. Pack, P. Israelsen, and K. Sealy, "A co-boresighted synchronized ladar/EO imager for creating 3D images of dynamic scenes," *Proc. SPIE* **5791**, 42–50 (2005).
4. R. T. Pack et al., "Eyesafe LADAR test-bed with coaxial color imager," *Proc. SPIE* **7323**, 732303 (2009).
5. B. M. Boldt et al., "A handheld texel camera for acquiring near-instantaneous 3D images," in *Conference Record of the Forty-First Asilomar Conference on Signals, Systems, and Computers, 2007. ACSSC 2007*, pp. 953–957, IEEE (2007).
6. M. Lindner, A. Kolb, and K. Hartmann, "Data-fusion of PMD-based distance-information and high-resolution RGB-images," in *Int. Symp. on Signals, Circuits and Systems*, Vol. 1, pp. 1–4, IEEE (2007).
7. M. Frank et al., "Theoretical and experimental error analysis of continuous-wave time-of-flight range cameras," *Opt. Eng.* **48**(1), 013602 (2009).
8. R. Ruocco, T. White, and D. Jarrett, "Systematic errors in active 3D vision sensors," *Proc. IEEE* **150**(6), 341–345 (2003).
9. M. Lindner et al., "Time-of-flight sensor calibration for accurate range sensing," *Comput. Vis. Image Underst.* **114**(12), 1318–1328 (2010).
10. I. Schiller, C. Beder, and R. Koch, "Calibration of a PMD camera using a planar calibration object together with a multi-camera setup," *Int. Arch. Photogramm. Remote Sens. Spatial Inform. Sci.* **XXXVII**(Part B3a), 297–302 (2008).
11. M. Lindner and A. Kolb, "Calibration of the intensity-related distance error of the PMD TOF-camera," *Proc. SPIE* **6764**, 67640W (2007).
12. T. Kahlmann, F. Remondino, and H. Ingensand, "Calibration for increased accuracy of the range imaging camera Swissranger," *Image Engineering and Vision Metrology* **36**(3), 136–141 (2006).
13. M. Lindner and A. Kolb, "Lateral and depth calibration of PMD-distance sensors," in *Int. Symp. on Visual Computing*, pp. 524–533, Springer-Verlag, Berlin- Heildelberg (2006).
14. M. Reynolds et al., "Capturing time-of-flight data with confidence," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 945–952, IEEE (2011).
15. M. Frank, M. Plaue, and F. A. Hamprecht, "Denoising of continuous-wave time-of-flight depth images using confidence measures," *Opt. Eng.* **48**(7), 077003 (2009).
16. S. E. Budge, "A calibration-and-error correction method for improved texel (fused ladar/digital camera) images," *Proc. SPIE* **8379**, 83790D (2012).
17. R. E. Altenhofen and R. T. Hedden, "Transformation and rectification," in *Manual of Photogrammetry*, M. M. Thompson, ed., Vol. 2, pp. 1–59, American Society of Photogrammetry, Falls Church, VA (1966).
18. K. Milburn, *Digital Photography: Expert Techniques*, O'Reilly & Associates Inc. (2003).
19. J. Heikkilä and O. Silvén, "Calibration procedure for short focal length off-the-shelf CCD cameras," in *Proc. of the 13th Int. Conf. on Pattern Recognition*, Vol. 1, pp. 166–170, IEEE (1996).
20. Y. Ma et al., *An Invitation to 3-D Vision: From Images to Geometric Models*, SpringerVerlag, New York City (2003).
21. Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *7th IEEE Int. Conf. Computer Vision*, Vol. 1, pp. 666–673, IEEE (1999).
22. J. Heikkilä and O. Silvén, "A four-step camera calibration procedure with implicit image correction," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1106–1112, IEEE (1997).
23. R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE Trans. Robot. Autom.* **3**(4), 323–344 (1987).
24. T. Melen, "Geometrical modeling and calibration of video cameras for underwater navigation," PhD thesis, Norges Tekniske Høgskole, Institutt for Teknisk Kybernetikk (1994).
25. C. Beder, B. Bartczak, and R. Koch, "A comparison of PMD-cameras and stereo-vision for the task of surface reconstruction using patchlets," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1–8, IEEE (2007).
26. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge University Press, Cambridge University Press, Cambridge, UK (2003).

27. J.-Y. Bouguet, "Camera calibration toolbox for Matlab," (May2012), http://www.vision.caltech.edu/bouguetj/calib_doc/.
28. A. M. Mayampurath, "Real-time cell image analysis," Master's thesis (Utah State University, Logan (2005).

Scott E. Budge received his BS, MS, and PhD degrees in electrical engineering from Brigham Young University in 1984, 1985, and 1990, respectively. He joined the faculty at Utah State University in January of 1989, where he has been involved in research into image data compression algorithms for transmission systems and space-based observation platforms. His current work involves research on methods for exploitation of full-waveform LADAR and electro-optic images for three-dimensional image applications. He is also involved in the development of high-performance image processing algorithms designed for implementation in hardware and VLSI systems.

Neeraj S. Badamikar received his bachelor of engineering degree in electronics and telecommunication in 2011 from Vishwakarma Institute of Information Technology, Pune, India. He is pursuing his master of science degree in electrical engineering from Utah State University, USA. His current research involves automatic registration of multiple texel images to form a combined texel image.