Linear Methods for Calibrating LIDAR-and-Camera Systems

Andrew Willis, Malcolm Zapata, James Conrad
University of North Carolina at Charlotte, Charlotte, NC 28223

Abstract*—This article describes a multimedia system consisting of two sensors: (1) a laser range scanner (LIDAR) and (2) a conventional digital camera. Our work specifies a mathematical calibration model that allows for this data to be explicitly integrated. Data integration is accomplished by calibrating the system, i.e., estimating for each variable of the model for a specific LIDAR-and-camera pair. Our approach requires detection of feature points in both the LIDAR scan and the digital images. Using correspondences between feature points, we can then estimate the model variables that specify an explicit mathematical relationship between sensed (x,y,z) LIDAR points and (x,y) digital image positions. Our system is designed for 3D line scanners, i.e., scanners that detect positions that lie in a 3D plane which requires some special theoretical and experimental treatment. Results are provided for simulations of the system in a virtual environment and for a real LIDAR-and-camera system consisting of a SICK LMS200 and an inexpensive USB web-camera. Calibrated systems can integrate the data in real-time which is of particular use for autonomous vehicular and robotic navigation.

Figure 1. An image of a calibrated LIDAR-and-camera system consisting of a SICK LMS200 and an inexpensive commodity USB web-camera.

I. Introduction

Camera calibration is a core computer vision problem that has been studied for over two decades. Seminal papers on the subject by Longuet-Higgins [1] and Tsai [2] and subsequent significant developments in [3], [4], [5] have brought the topic of camera calibration to some level of maturity (see [6] for a recent review of calibration techniques). Others researchers formulate the camera calibration problem in terms of a homography between the camera pair which incorporates both extrinsic and intrinsic parameters. The homography may be estimated using the point correspondence described in [2], [3], [4] or by finding vanishing points in the image by detecting the intersection points parallel 3D lines when projected into each camera image [5], [7]. This paper adopts time-proven concepts from the literature (specifically concepts from [2]) to develop a new method capable of calibrating a system consisting of a LIDAR line-scanner and a digital camera. Calibration of these two sensors allows their recorded data to be easily integrated. We have found no published methods to solve this problem which, with the availability of inexpensive 3D line-scanners, promises to be an area of need in the future.

II. Modeling and Analytical Derivation

Our method for calibrating a LIDAR-and-camera system is modeled after a classical method originally proposed by Tsai [2] that has been subsequently improved by a number of researchers. This calibration technique assumes a mathematical model for the camera that specifies how 3D (x, y, z) points are projected to (x, y) pixel locations within a camera's 2D image. Let $\mathbf{P}_w = (X_w, Y_w, Z_w)^t$ denote points measured in 3D by the LIDAR sensor and $\mathbf{p}_c = (x_{im}, y_{im})^t$ denote pixels locations in the 2D camera image. As in [2], we use a perspective camera model which, for a suitable choice of coordinate system, has the following projection equations:

$$x = f\frac{X_w}{Z_w} \ , \ \ y = f\frac{Y_w}{Z_w} \tag{1}$$

where f from equation (3) is the focal length of the camera in mm, and (x, y) from equation (3) is the physical (x, y) image location in mm corresponding to the projection of the 3D point (X_w, Y_w, Z_w). We denote the image pixel containing the physical location (x, y) as (x_{im}, y_{im}). Their values may be computed as follows:

$$\begin{aligned} x &= (x_{im} - o_x)s_x \\ y &= (y_{im} - o_y)s_y \end{aligned} \tag{2}$$

where (o_x, o_y) denotes the (x, y) position of the image center (in pixels) and (s_x, s_y) denotes the physical size of a single pixel in the image sensor and has units $mm/pixel$. Our calibration procedure uses equations (3) and (2) to relate 3D points, $\mathbf{P}_w$, measured by the LIDAR sensor, to 2D image pixels, $\mathbf{p}_c$. When combined, the resulting equations form a linear system of equations.

Our calibration method uses a geometric calibration pattern consisting of a sequence of rectangular boxes that are placed adjacent to one-another such that at least two faces are visible in the camera calibration images. The faces of each box are colored such that adjoining box faces alternate between

black and white. The calibration process proceeds by obtaining measurements of the calibration pattern surface with both the LIDAR and camera sensors. For the LIDAR sensor, a sequence of 3D points is obtained that, for convenience, is assumed to lie in the $xz-plane$, i.e., the 3D plane perpendicular to the $y-axis$ that contains the origin. Hence, all measured LIDAR data points have $Y_w = 0$ and the world coordinate system origin corresponds to the measurement reference point of the LIDAR sensor. For the camera sensor, an image is recorded that also includes intensity measurements of the calibration pattern surface.

We then process the recorded data to detect a sequence of feature locations. The specific features we seek are locations where the calibration box faces touch. In the camera image, these locations are found by detecting locations where the intensity changes between black and white which are found via standard image edge detection techniques, e.g., the Canny edge detector. In the LIDAR data, these locations are found via a two-step process: (1) a sequence of lines are fit to the sensed 3D data and (2) the intersection points of the fit lines correspond to 3D locations where the box faces touch.

Each box is oriented such that several (> 5) measurements are obtained by each sensor from two faces of each box. This generates three feature locations for the first calibration box placed in the scene and, since adjacent boxes share a feature location, two additional feature locations may be captured from each additional box in the calibration pattern.

Let $\mathbf{P}_{w,i}$ denote the i^{th} detected LIDAR data feature location listed in order of increasing value for the X_w coordinate. Let $\mathbf{p}_{c,j}$ denote the j^{th} detected camera image feature location listed in order of increasing value for the x_{im} coordinate. For calibration, we seek to find corresponding pairs of feature locations from the measured LIDAR-and-camera data. Corresponding feature pairs are denoted by their (i,j) indices and specify that the i^{th} LIDAR feature location, $\mathbf{P}_{w,i}$, is a measurement of the same surface location as the j^{th} image feature location, $\mathbf{p}_{c,j}$. Our calibration method expresses the unknown calibration parameters as a linear function of the estimated corresponding feature locations. However, since the LIDAR and camera sensors measure the surface in different coordinate systems, the geometric relationship between the two sensors must also be estimated. When the camera is moved away from the origin and it's orientation is modified the projection equations (3) and (2) change as follows:

$$x_{im} = \frac{f}{s_x}\frac{\mathbf{r}_1^t(\mathbf{P}_w - \mathbf{T})}{\mathbf{r}_3^t(\mathbf{P}_w - \mathbf{T})} + o_x \ , \ \ y_{im} = \frac{f}{s_y}\frac{\mathbf{r}_2^t(\mathbf{P}_w - \mathbf{T})}{\mathbf{r}_3^t(\mathbf{P}_w - \mathbf{T})} + o_y \tag{3}$$

where $\mathbf{r}_i^t$ denotes the 3x1 vector formed by the i^{th} row of the rotation matrix that determines the camera orientation and $\mathbf{T} = (X, Y, Z)$ denotes the 3D position of the camera optical center relative to the LIDAR system measurement center. For LIDAR line scanners, there is a difficulty in calibration due to the fact that the scanner data is constrained to lie in the $xz - plane$. Hence the data does not sufficiently constrain the position and orientation of the camera to unambiguously solve for the complete 3D camera position and orientation. For this reason, we assume that the y-component of the camera position is *a-priori* known, i.e., $T_y = Y_0$, and the camera's optical axis is parallel to the $xz - plane$. This implies that $\mathbf{r}_2^t = (0, 1, 0)$, i.e., that the camera orientation can only change in the $xz - plane$, and that we can only estimate the X and Z components of $\mathbf{T}$, the camera position relative to the LIDAR sensor. These assumptions serve to simplify the stated projection equations as shown below:

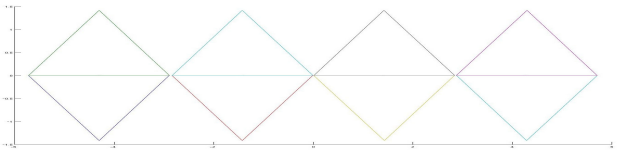

Figure 2. A top-down view of the 3D model used as the simulated calibration pattern. It consists of four adjacent boxes that touch at their corners.

$$y_{im} = f_y\frac{Y_0}{r_{31}X_w + r_{33}Z_w - T_z} \ , \ x_{im} = f_x\frac{r_{11}X_w + r_{13}Z_w - T_x}{r_{31}X_w + r_{33}Z_w - T_z} \tag{4}$$

where we have merged products of unknown parameters into newly defined unknowns as follows: $T_x = \mathbf{r}_1^t\mathbf{T}$, $T_z = \mathbf{r}_3^t\mathbf{T}$, $f_x = \frac{f}{s_x}$, $f_y = \frac{f}{s_y}$. We can then write the linear equations of the LIDAR-and-camera system by placing all terms of equations (4) on the left hand side:

$$\begin{gathered} r_{31}x_{im}X_w + r_{33}x_{im}Z_w - x_{im}T_z - f_x r_{11}X_w - f_x r_{13}Z_w + f_x T_x = 0 \\ r_{31}y_{im}X_w + r_{33}y_{im}Z_w - y_{im}T_z - f_y Y_0 = 0 \end{gathered} \tag{5}$$

We then use the estimated correspondence pairs $(\mathbf{P}_{w,i}, \mathbf{p}_{c,j})$ to construct a single equation in the linear system (5) whose purpose is to estimate the unknown calibration parameters. In practice, one cannot know the exact row position of the LIDAR line scanner in the image (the value of y_{im} for the second equation) since LIDAR scanners typically use infra-red laser diodes that illuminate the scene surfaces at a wavelength that is greatly attenuated (or undetectable) by the image sensor. Pragmatically, this prevents use of the second equation from (5) for calibration in our applications since it involves the unknown value of y_{im} for each detected feature location. However, one should note that this analysis raises the possibility of being able to directly obtain highly accurate LIDAR-and-camera calibrations via infra-red sensitive cameras, a subject not treated in the scope of this article.

III. Overview of Approach

Our calibration approach is implemented via the following sequence of steps:

1) Fix the camera and LIDAR sensor into a relative geometry where the camera optical axis is parallel to the measurement plane of the scanner (see Figure 1).
2) Prepare a calibration target consisting of adjacent (touching) boxes with different colors on each box face (see Figure 2). Place the boxes in front of the LIDAR-and-camera sensor and obtain measurements of the calibration pattern surface from each sensor. Boxes should be standing on a surface parallel to the measurement planes of the LIDAR scanner and camera such that the vertical box-edges appear as vertical lines in the recorded camera images (see Figure 3). *Note: Calibration will still be successful if there are deviations from this setup but these deviations may cause distortion in the estimated calibration parameters.*

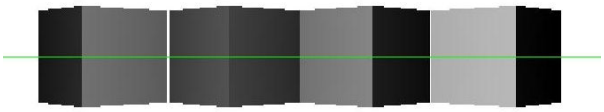

Figure 3. A frontal view of the simulated virtual geometric calibration pattern is shown. The generated image is the image recorded by a virtual camera that serves to simulate the proposed calibration technique. A green horizontal line across the image denotes the projected location of the LIDAR sensor $xz-plane$ as it appears from the view of the virtual camera.

3) Estimate feature locations in the sensed camera image, i.e., pixel column locations $x_{im,j}$ where the box face colors change, and the world coordinate system x locations where the box faces touch, i.e., the 3D x-positions of the box corner locations, $X_{w,i}$.
4) Match up the feature locations so that the (i,j) pairs $(x_{im,j}, X_{w,i})$ is the pixel location of a box corner and the 3D position of the same box corner respectively. There must be at least 6 such pairs to generate a unique solution for the six unknowns in the top equation of the linear system described by equation (5).
5) Each (i,j) pair provides the needed values for a row of the matrix $\mathbf{A}$ shown below. Again, there must be at least 6 such pairs. Additional pairs serve to improve the accuracy of the estimated calibration parameters.

$$\mathbf{A} = \begin{bmatrix} x_{im}X_w & x_{im}Z_w & -x_{im} & -X_w & -Z_w & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{im}X_w & x_{im}Z_w & -x_{im} & -X_w & -Z_w & 1 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} r_{31}, & r_{33}, & T_z, & f_x r_{11}, & f_x r_{13}, & f_x T_x \end{bmatrix}^t$$

6) Solve the system of equations $\mathbf{Ax} = 0$ via singular value decomposition (SVD). SVD will decompose the matrix $\mathbf{A}$ into the matrix product $\mathbf{A} = \mathbf{UDV}^t$. We then find the index i of the minimum singular value along the diagonal of the matrix $\mathbf{D}$ and form the column vector $\mathbf{v}_i$ as the elements of column i in the matrix $\mathbf{V}$. $\mathbf{v}_i$ is the solution vector; a vector corresponding to the least-squares error estimate of the values in the vector $\mathbf{x}$ from step 5. Since SVD automatically normalizes the columns of $\mathbf{V}$ such that $\|\mathbf{v}_i\| = 1$, this vector is multiplied by an unknown scale factor, γ.
7) Using knowledge that the correct rotation vector $\mathbf{r}_3^t = (r_{31}, 0, r_{33})$ is constrained to have unit length, we can compute the value of this scale factor from the first two components of the vector $\mathbf{v}_i$: $\gamma = \sqrt{v_1^2 + v_2^2}$. Formulas for the other parameters of the calibration are: $\mathbf{r}_3^t = (\frac{v_1}{\gamma}, 0, \frac{v_2}{\gamma})$, $T_z = \frac{v_3}{\gamma}$, $f_x = \sqrt{\frac{v_4^2+v_5^2}{\gamma^2}}$, $\mathbf{r}_1^t = (\frac{v_4}{f_x\gamma}, 0, \frac{v_5}{f_x\gamma})$, $T_x = \frac{v_6}{f_x\gamma}$ and $\mathbf{r}_2 = \mathbf{r}_3 \times \mathbf{r}_1 = (0,1,0)$. These variables determine the calibration parameters of the LIDAR-and-camera sensor system.

IV. Simulation, Results, and Conclusion

The proposed calibration method was simulated in a virtual 3D environment (see Figures 2, 3, 4) and then implemented using a SICK LMS200 LIDAR sensor and an inexpensive commodity web-camera (see Figure 1). The simulation of the system applied the mathematical models described from §II to generate a simulated camera image (see Figure 3) and a simulated LIDAR scan (see Figure 4) to validate, debug and perform stability analysis on the proposed calibration method with acceptable results. Afterward, the real-world LIDAR-and-camera system shown in Figure 1 was calibrated and used to generate the image shown in Figure 5. Note that some calibration error is evident but there is good agreement between the LIDAR range data and the (x,y) locations of the associated pixels. This is true for large scene depth variations as exhibited by the significant change in the row positions of the projected LIDAR points for trees in the foreground and background in Figure 5.

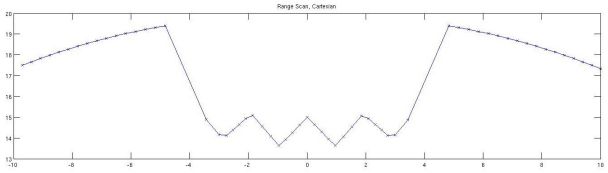

Figure 4. A plot of the sensed LIDAR depths observed from the simulated virtual calibration pattern from Figure XX. Depth measurements are divided into local linear segments by line-fitting and the calibration pattern corners are extracted.

Figure 5. An example of a scene exhibiting large depth variations. The 3D locations of the LIDAR-sensed depth data has been projected into the camera image and is shown as a sequence of blue points within the image.

References

[1] H. C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature*, no. 293, pp. 133–135, 1981.
[2] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras," *IEEE Journal Robotics and Automation*, vol. 3, no. 4, pp. 323–344, 1987.
[3] Heikkil and Silven, "A four-step camera calibration procedure with implicit image correction," in *CVPR*, pp. 1106–1112, 1997.
[4] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1130–1134, 2000.
[5] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
[6] F. Remondino and C. Fraser, "Digital camera calibration methods: considerations and comparisons," in *IEVM06*, pp. 266–272, 2006.
[7] L. Grammatikopoulos, G. Karras, and E. Petsa, "An automatic approach for camera calibration from vanishing points," *ISPRS journal of photogrammetry and remote sensing*, vol. 62, no. 1, pp. 64–76, 2007.