Calibration of Dual Laser-Based Range Cameras for Reduced Occlusion in 3D Imaging

Aaron Mavrinac, Xiang Chen, Peter Denzinger, and Michael Sirizzotti

Abstract— A robust model-based calibration method for dual laser line active triangulation range cameras, with the goal of reducing camera occlusion via data fusion, is presented. The algorithm is split into two stages: line-based estimation of the lens distortion parameters in the individual cameras, and computation of the perspective transformation from each image to a common world frame in the laser plane using correspondences on a target with known geometry. Experimental results are presented, evaluating the accuracy of the calibration based on mean position error as well as the ability of the system to reduce camera occlusion.

I. Introduction

Active 3D vision is a popular family of methods for obtaining robust and accurate three-dimensional digitizations of real objects [1], [2]. One common paradigm, known as laser line triangulation, uses a single camera to view a laser line projected over the surface of an object at some angular offset from its principal axis, processing the image to obtain a cross-section profile of the object; this process can be repeated moving the object in small increments normal to the laser plane to yield a full surface scan of the object.

A major problem with these devices, particularly in inspection and metrology applications, is the occlusion of portions of the target object, resulting in incomplete 3D data. Two types of occlusion may occur: *laser occlusion*, in which the laser is unable to illuminate an object point visible from the camera, and *camera occlusion*, in which the camera is unable to image an object point illuminated by the laser. Such issues are typically overcome by performing multiple scans or employing more complex systems, imposing a large amount of overhead.

Camera occlusion occurs in any case where a portion of the target surface faces away from the camera at a greater angle from the horizontal than the camera itself (about the x axis, in our convention described in Section II-A). Thus, it is possible to mitigate this by adding a second camera to the system at some angle on the opposite side of the laser plane. In practice, most surface portions occluded from one direction are visible from the other, with the occlusion typically being caused by a height discontinuity of some sort; eliminating this therefore usually yields nearly complete 3D

data. We seek to obtain a combined range image with all information which would otherwise be available from two separate scans in opposite orientations.

The challenge, then, is to combine the data from both sources in such a way that existing processes (e.g. for inspection or metrology) can be applied to the more complete data unmodified – in other words, the system should be a drop-in replacement for the single-camera equivalent. We present here a robust model-based calibration method for two laser-based range cameras which allows this data to be fused into a single 3D point cloud or range image in real time.

Although model-based calibration of active triangulation camera systems is essentially the same problem as standard camera calibration, some techniques taking advantage of the specifics of laser line triangulation have been proposed. Reid [3] presents a method for estimating the projective homography with the laser plane using correspondences between the image and a set of orthogonal planes of known geometry in the scene. Jokinen [4] presents an area-based matching approach in which multiple profile maps from different viewpoints are registered to refine an initial target-based calibration.

Departing from model-based calibration, Trucco et al. [5] present a direct calibration method which interpolates a lookup table for the entire field of view based on a target of known geometry, and thus implicitly models all intermediate parameters; this is further explored in [6].

Vilaça et al. [7] present a complete calibration method for two laser-based range cameras, also with the goal in mind of reducing camera occlusion. It is similar to our method in that it constrains lens distortion correction and the perspective homography to the laser plane – a valid simplification over traditional camera calibration, since subsequent measurements are also constrained there. However, our approach uses the range data directly for calibration, which allows for implicit constraint of calibration to the laser plane, higher accuracy (if range values can be obtained with subpixel accuracy), a more direct line-based process for lens distortion correction, and the use of a more practical calibration apparatus.

In the general case, combining range data from multiple sources is often achieved via registration algorithms (Salvi et al. [8] present an excellent overview). While this approach is well-studied and its various algorithms can be applied in a diverse range of situations, the calibration approach has clear advantages: it is completely unaffected by incomplete overlap, which in contrast causes severe performance degradation

or imposes additional overhead in registration algorithms; also, a pre-computed lookup table is far less computationally expensive than iterative registration, and allows for more exact results. Laser-based range cameras conveniently lend themselves to such calibration.

The remainder of this paper is organized thus. Section II introduces some concepts, conventions, and notation used subsequently. The proposed calibration method is detailed in Section III. Experimental results are presented in Section IV. Finally, some concluding remarks are given in Section V.

II. DEFINITIONS

A. Geometry

The plane through which the laser line is projected is defined as the $x - z$ plane in the world coordinate system (with x horizontal and z vertical), and is termed the *laser plane*. The direction a target object moves when performing a scan is termed the *transport direction*; this must, of course, have a y component. World coordinates are assumed to be defined in some real measurement unit; we use millimetres herein.

We assume a roughly symmetric camera configuration (Figure 1), in which the transport direction is positive-y (normal to the laser plane), and the cameras are placed on opposite sides of the laser plane. Other configurations are possible, and our method can be applied to these as well with minimal modification.
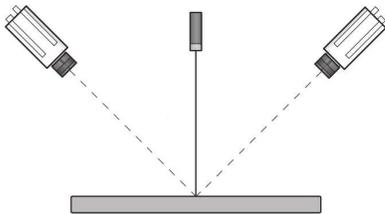

Fig. 1.   Symmetric Camera Configuration

The raw discrete two-dimensional coordinates of the data from the camera, denoted (u_r, v_r), lie in the *sensor plane*. A continuous *image plane*, with coordinates denoted (u, v), is defined to describe corrected (or ideal) data points.

B. Range Data

We assume that the cameras in question are already able to perform basic laser line triangulation, either on-board or in software on a host computer, using well-studied methods (see Section IV-A for the specifics of our experimental setup and imaging hardware).

The range data generated by the camera from a single image is termed a *profile*, and consists of an ordered set of height values – one per camera sensor column – corresponding to the z values of object points in the laser plane. Profile elements are points in the sensor plane, with column index u_r and height value v_r (height values may be interpolated and are thus not necessarily actual sensor row indices). An ordered set of profiles, ostensibly taken at regular intervals along the transport direction, is termed a *scan*.
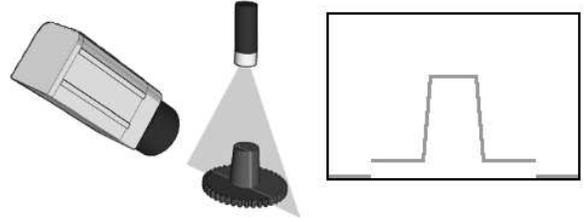

Fig. 2.   Range Data

III. CALIBRATION METHOD

Our calibration method is divided into two stages: correction for lens distortion and determination of a perspective mapping from the image plane to the laser plane.

A. Lens Distortion

According to Brown's model of lens distortion [9], image plane coordinates (u, v) are computed from sensor plane (raw pixel) coordinates (u_r, v_r) as follows:

$$u = u_r + u_0(K_1 r^2 + K_2 r^4 + \ldots) + P_1(r^2 + 2u_0^2) + 2P_2 u_0 v_0(1 + P_3 r^2 + \ldots) \tag{1}$$

$$v = v_r + v_0(K_1 r^2 + K_2 r^4 + \ldots) + P_2(r^2 + 2v_0^2) + 2P_1 u_0 v_0(1 + P_3 r^2 + \ldots) \tag{2}$$

$$u_0 = u_r - o_u \qquad v_0 = v_r - o_v \qquad r = \sqrt{u_0^2 + v_0^2}$$

where (o_u, o_v) is the optical center on the sensor plane, K_i are radial distortion coefficients, and P_i are tangential (or decentering) distortion coefficients. For practical purposes, we limit our model to coefficients K_1, K_2, P_1, and P_2; experiments have shown that higher-order coefficients are relatively insignificant in most cases [10].

The parameters can be estimated by exploiting the well-known fact that, in the absence of distortion, straight lines in the three-dimensional scene map to straight lines in the two-dimensional image [11], [12]. Line-based correction lends itself particularly well to our case, since we have a readily-available source of straight lines in the scene constrained to the plane of interest.

1) Line Extraction: Line point sets are obtained by taking a profile of any flat object. The raw profile may contain extraneous points; these are eliminated using a two-step process: first, all points with height below a certain threshold are removed; second, the remaining points are fit to a linear model using RANSAC [13]. The linear model takes the form:

$$v = \alpha u + \beta \tag{3}$$

with the origin (o_u, o_v) at the optical center (using the sensor center as an initial guess). Since, clearly, we cannot yet obtain u and v, we substitute the raw values u_r and v_r, respectively, assuming that lens distortion is negligible for the purpose of line extraction. The RANSAC consensus set

forms the final line point set, and the model parameters α and β are retained to initialize the optimization model in the following step.

A number of lines, well-distributed over the field of view, should be profiled. Although not necessary, it is convenient to do so with both cameras simultaneously.

2) Parameter Optimization: Given a set of M line point sets, each composed of N_m points, the objective is to find the distortion parameters such that the total deviation in the line equations is minimized [11]. This is achieved by minimizing F, the sum of squared distances from each undistorted point (u, v) to its corresponding line:

$$f = \alpha_m u - v + \beta_m \qquad (4)$$

$$F = \sum_{m=1}^{M} \sum_{k=1}^{N_m} f^2 \qquad (5)$$

where u and v are computed according to Equations 1 and 2, respectively. The set of $2M + 6$ design variables is $\{K_1, K_2, P_1, P_2, o_u, o_v, \alpha_1 \dots \alpha_M, \beta_1 \dots \beta_M\}$. The distortion parameters (K_i and P_i) are initialized to zero, (o_u, o_v) is initialized to the sensor center, and the line parameters (α_m and β_m) are initialized from the RANSAC model parameters from the extraction step.

A solution can be found numerically using the Levenberg-Marquardt algorithm [14] for nonlinear optimization (the authors of [11] use a simpler gradient descent method, but note problems with convergence to local minima). This process requires the Jacobian matrix, consisting of the partial derivatives of f with respect to each parameter.

B. Perspective Mapping

1) Homography: A homography between the two-dimensional image plane and the two-dimensional laser plane is defined by a 3×3 matrix $\mathbf{H}$ as:

$$\lambda \begin{bmatrix} x \\ z \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \qquad (6)$$

where λ is a scale factor.

Given a set of laser plane point coordinates and their corresponding image plane coordinates, this homography can be solved linearly. A point correspondence pair $(x, z) \leftrightarrow (u, v)$ yields two linear equations:

$$x(h_{31}u + h_{32}v + h_{33}) - (h_{11}u + h_{12}v + h_{13}) = 0 \quad (7)$$

$$z(h_{31}u + h_{32}v + h_{33}) - (h_{21}u + h_{22}v + h_{23}) = 0 \quad (8)$$

At least five such point correspondences are required to find $\mathbf{H}$. An optimal solution to the resulting overdetermined linear system can be found numerically using singular value decomposition.

2) Calibration Target: Point correspondences are obtained by taking a single profile of a calibration target of known structure. The target should have a number (five or more) of precise, well-distributed, non-collinear points, such as sharp corners, which can be easily localized in the image (see Section IV-A.2 for our design and a brief description of our point detection method). The detected image plane point (u, v) is associated with the known laser plane point (x, z) defined in units of length relative to a common origin.

The profile is taken by both cameras simultaneously, so that their independent homographies map to a common coordinate system.

IV. EXPERIMENTAL RESULTS

A. Apparatus

1) Camera System: All experiments were performed using two SICK-IVP Ranger D industrial 3D smart cameras with a single laser line source. These were mounted in a fixed configuration similar to Figure 1 above a small conveyor belt, with an encoder and photoswitch for controlling the start and transport direction resolution of scans. Figure 3 shows the equipment during the calibration procedure.
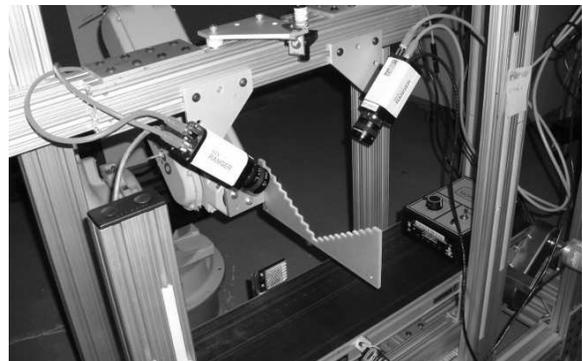

Fig. 3. Equipment Configuration

2) Calibration Target: We employ a calibration target with one flat side (for the lens distortion correction stage) and one side with a number of right-angled steps (for the perspective mapping stage), as seen in Figure 3. Our target was manufactured to a tolerance of 0.1 mm. This is attached to a robotic arm to automate the positioning of the target. Although accurate positioning is not required, rotation of the target about the y and z world coordinate axes introduces error and should be avoided.

Nine line profiles on the flat side are taken for use in the lens distortion step, in a pattern spanning the field of view. Figure 4 shows a composite of data from all nine scans for one camera; note the significant radial distortion prior to correction.

Point detection on the stepped side is performed by finding the middle (lowest) step, then searching outwards for the horizontal lines of each step, using RANSAC to extrapolate the precise corner positions. Figure 5 shows a plot of a typical set of results, with squares indicating the estimated corner position.
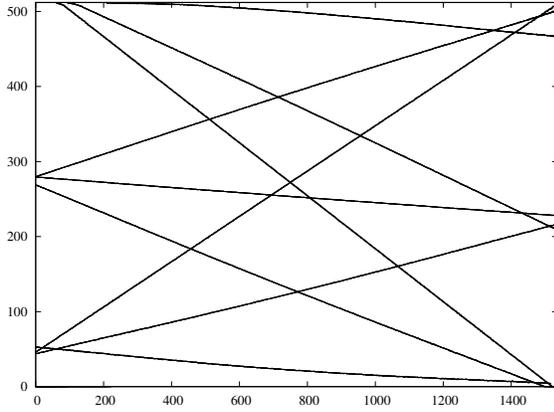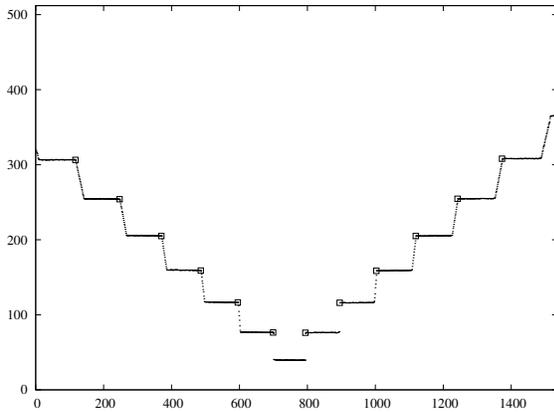
Fig. 4.   Line Profile Pattern

Fig. 5.   Typical Point Detection Result

3) Software: The software implementation is developed in ANSI C and compiled with GCC using the MinGW tools. The SICK iCon API is used to interface with the range cameras over Gigabit Ethernet. For Levenberg-Marquardt optimization, singular value decomposition, and random number generation for RANSAC, the GNU Scientific Library [15] is employed. Range images are output in PGM (P2 ASCII) format.

B. Accuracy

The accuracy of the sensor calibration is tested by measuring a large number of points of known geometry relative to a common origin. The mean error is computed separately for each camera, yielding the calibration accuracy for measurements on each side, and then the mean deviation for each point from one camera to the other is computed, yielding the relative accuracy of the pair.

The sensor in the Ranger D camera has 1536 columns and 512 rows. To interpolate the impact position of the laser line on the sensor, a built-in algorithm using a center-of-gravity method with a typical resolution of 1/10 pixel is employed; results are returned in terms of 1/16 pixel, yielding an effective sensor size of 1536×8192. Our calibration software

generates a lookup table between these sensor coordinates and world coordinates in the laser plane, by mapping each discrete coordinate through the distortion model and the perspective homography.

For our test case with a field of view 150 mm high by 200 mm wide and the cameras at an angle of $45°$ to the laser plane, the mean absolute position errors for points in the individual cameras were measured at 0.1007 mm for the left and 0.0968 mm for the right (on the same order as the manufacturing tolerance for our calibration target), while the mean relative position error between the two cameras was measured at 0.2123 mm (note that the cameras estimate the laser plane origin independently, hence the larger relative error). This corresponds to subpixel accuracy in our output images, which are sampled at 0.5 mm per pixel. Adjusting for the relative size of field of view and camera resolution, these results compare favorably with those of Vilaça et al. [7].

C. Occlusion Reduction

Our measurement software generates one row of a range image by binning and averaging the points in a rectified profile – generated by combining the raw profiles from each camera transformed through their respective lookup tables – according to a given field of view width and resolution. Where no data is available, a value of zero is assigned. A full range image is generated from a scan, with the aforementioned resolution typically equal to the spacing between profiles in the scan.

As an approximate quantification of camera occlusion, we define a closed countour of the scanned object in the range image by segmentation, and consider the proportion of pixels within this contour with value zero (black pixels) expessed as a percentage. We compare the occlusion in the left and right range images (that is, the range images generated using only data from the left and right cameras, respectively) to that in the combined image for a number of objects, with results shown in Table I.

TABLE I
OCCLUSION REDUCTION FOR VARIOUS OBJECTS

Object	Occl. Left	Occl. Right	Occl. Combined
Connecting Rod	8.35%	8.20%	1.22%
Pump Inserts	3.64%	3.80%	0.11%
Transmission Gear	3.10%	3.28%	0.10%
Subfloor Panel	8.37%	14.91%	0.82%
Bearing Collar	16.48%	14.62%	2.85%
Toy Bricks	17.14%	13.49%	2.49%

Figures 6 and 7 show the calibrated images from the left and right cameras as well as the combined image for two typical objects. The effect of camera occlusion can clearly be seen in the left and right images: black regions near raised edges show where no data exists. By contrast, these effects are almost completely absent from the combined images.

These results highlight the problem faced in many applications as well as the impact of the solution. Regardless of how objects are oriented, with single-camera scans, there

is in the general case some degree of camera occlusion. When the images from two cameras are combined, this occlusion is dramatically reduced (with minimal computation as compared to registration).
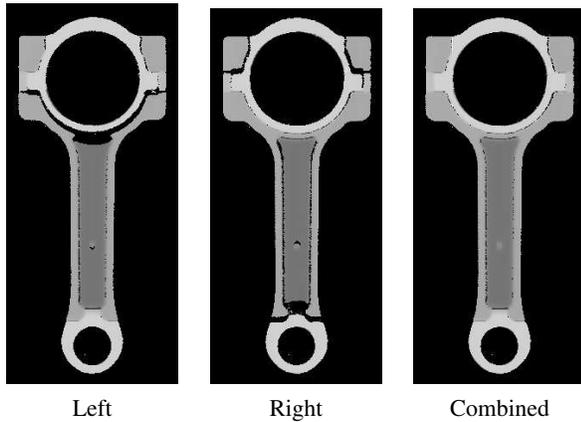

| Left | Right | Combined |

Fig. 6.   Occlusion Reduction for Connecting Rod
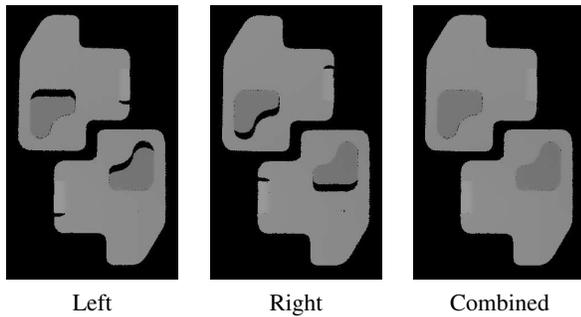

| Left | Right | Combined |

Fig. 7.   Occlusion Reduction for Pump Inserts

V. CONCLUSIONS

We have presented a straightforward calibration method for laser-based range cameras, which adapts the underlying theory of proven techniques to take advantage of the planar constraint and readily available straight line data inherent to these systems. For the individual sensor, the calibration has been shown to exhibit high absolute accuracy. The method is used to calibrate dual cameras into a common reference frame, with which we have achieved a substantial reduction in camera occlusion with high relative accuracy.

Although we present only dual cameras in a symmetric configuration here, the principles could easily be extended to more cameras and different configurations, potentially resulting in further reduction of occlusion for some applications. The calibration method can also be applied to single-camera systems.

Our method is applicable to a variety of inspection, metrology, and other tasks which might benefit from more complete range data in a single scan. Importantly, it allows existing methods working with range images to function unchanged with the fused data, making the system suitable as a drop-in replacement for a single-camera solution.

REFERENCES

[1] P. J. Besl, "Active, Optical Range Imaging Sensors," *Machine Vision and Applications*, pp. 127–152, 1988.
[2] S. K. Mada, M. L. Smith, L. N. Smith, and S. Midha, "An Overview of Passive and Active Vision Techniques for Hand-Held 3D Data Acquisition," in *Opto-Ireland 2002: Optical Metrology, Imaging, and Machine Vision*.   SPIE, 2003, pp. 16–27.
[3] I. D. Reid, "Projective Calibration of a Laser-Stripe Range Finder," *Image and Vision Computing*, vol. 14, no. 9, pp. 659–666, 1996.
[4] O. Jokinen, "Self-Calibration of a Light Striping System by Matching Multiple 3-D Profile Maps," in *Proc. 2nd Intl. Conf. on 3-D Digital Imaging and Modeling*, 1999, pp. 180–190.
[5] E. Trucco, R. B. Fisher, and A. W. Fitzgibbon, "Direct Calibration and Data Consistency in 3-D Laser Scanning," in *Proc. British Machine Vision Conf.*, 1994, pp. 489–498.
[6] E. Trucco, R. B. Fisher, A. W. Fitzgibbon, and D. K. Naidu, "Calibration, Data Consistency and Model Acquisition with a 3-D Laser Striper," *Intl. Jrnl. of Computer Integrated Manufacturing*, vol. 11, no. 4, pp. 292–310, 1998.
[7] J. L. Vilaça, J. C. Fonseca, and A. M. Pinho, "Calibration Procedure for 3D Measurement Systems Using Two Cameras and a Laser Line," *Optics & Laser Technology*, vol. 41, no. 2, pp. 112–119, 2009.
[8] J. Salvi, C. Matabosch, D. Fofi, and J. Forest, "A Review of Recent Range Image Registration Methods with Accuracy Evaluation," *Image and Vision Computing*, vol. 25, pp. 578–596, 2007.
[9] D. C. Brown, "Decentering Distortion of Lenses," *Photogrammetric Engineering*, vol. 32, no. 3, pp. 444–462, 1966.
[10] J. Weng, P. Cohen, and M. Herniou, "Camera Calibration with Distortion Models and Accuracy Evaluation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, no. 10, pp. 965–980, 1992.
[11] B. Prescott and G. F. McLean, "Line-Based Correction of Radial Lens Distortion," *Graphical Models and Image Processing*, vol. 59, no. 1, pp. 39–47, 1997.
[12] F. Devernay and O. Faugeras, "Straight Lines Have to be Straight," *Machine Vision and Applications*, vol. 13, pp. 14–24, 2001.
[13] M. A. Fischler and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Comm. of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
[14] C. T. Kelley, *Iterative Methods for Optimization*.   SIAM, 1999.
[15] M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, P. Alken, M. Booth, and F. Rossi, *GNU Scientific Library Reference Manual*, 3rd ed. [Online]. Available: http://www.gnu.org/software/gsl/