Extrinsic Calibration of a Camera and Laser Range Finder (improves camera calibration)

Qilong Zhang
Department of Computer Science and Engineering
Washington University in St. Louis
St. Louis, MO. 63130 United States
Email: zql@cse.wustl.edu

Robert Pless
Department of Computer Science and Engineering
Washington University in St. Louis
St. Louis, MO. 63130 United States
Email: pless@cse.wustl.edu

Abstract— We describe theoretical and experimental results for the extrinsic calibration of sensor platform consisting of a camera and a 2D laser range finder. The calibration is based on observing a planar checkerboard pattern and solving for constraints between the "views" of a planar checkerboard calibration pattern from a camera and laser range finder. we give a direct solution that minimizes an algebraic error from this constraint, and subsequent nonlinear refinement minimizes a re-projection error. To our knowledge, this is the first published calibration tool for this problem. Additionally we show how this constraint can reduce the variance in estimating intrinsic camera parameters.

I. INTRODUCTION

In the recent years, two dimensional laser ranger finders mounted on mobile robots have become very common for various robot navigation tasks. They provide in real time accurate range measurements in large angular fields at a fixed height above the ground plane, and enable robots to perform more confidently a wide range of tasks by fusing image data from the camera mounted on robots [12], [1], [5], [9]. In order to effectively use the data from the camera and laser range finder, it is important to know their relative position and orientation from each other, which affects the geometric interpretation of its measurements.

The calibration of each of these geometric sensors can be decomposed into internal calibration parameters and external parameters. The external calibration parameters are the position and orientation of the sensor relative to some fiducial coordinate system. The internal parameters, such as the calibration matrix of a camera, affect how the sensor samples the scene. This work assumes the internal sensor calibration is known, and focuses on the external calibration. Here we propose a method for extrinsic calibration of a camera and laser range finder, that is, identifying the rigid transformation from the camera coordinate system to the laser coordinate system. The method employs a planar calibration pattern viewed simultaneously by the camera and laser range finder. For each different pose of the planar pattern, the method constrains the extrinsic parameters by registering the laser scanline on the planar pattern with the estimated pattern plane from the camera image.

It is important also to differentiate this work from the problems that at first may appear similar. There has been a great deal of work on calibration for laser scanners, which
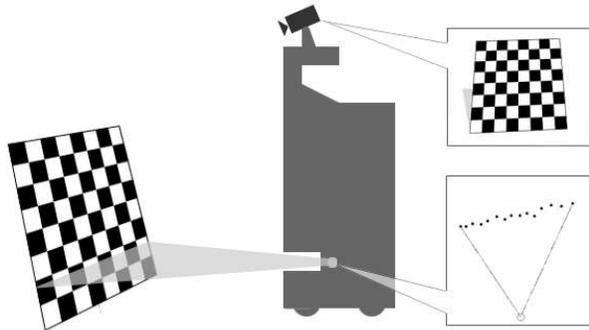

Fig. 1. A schematic of the calibration problem considered here. A planar calibration pattern is posed in the both views of the camera and the laser range finder.The goal of this paper is to study a calibration method that finds the rotation Φ and the translation Δ which transform points in the camera coordinate system to points in the laser coordinate system

are the parts of active vision systems that project a point or a stripe which is then viewed by the camera. Finding the geometric relationship between the laser scanner and the camera is vital to creating metric depth estimates to build textured 3D models, for example [3]. Calibration methods exist for this problem, which make use of the visible position of the laser point or stripe [7]. In this paper we consider an extrinsic calibration of a camera with a laser range finder where the laser points are invisible to the camera. This calibration applies to a very common sensor package for a large number of autonomous robots, such as the iRobot series, and there is no calibration method published to date.

Even though there has been increasing use of 3D laser range finders, they are still lack of portability and flexibility. Furthermore, the time cost of 3D data acquisition is also very expensive, since the systems require time to scan the laser through different directions in the environment. For many robotic tasks, such as robot navigation, it may be more important to scan over a smaller area at a higher frequency, which allows autonomous robots sense the environment in real time and to act on the basis of acquired data. That's why we focus on the pose estimation of camera w.r.t 2D laser range finder, which is cost-effective while provides flexibility and accuracy for range data acquisition.

The remainder of this paper is organized as follows. The next section introduces the basic equations associated with the extrinsic calibration, formalizes the problem we are going to solve and derives the geometric constraints on the rigid transformation from a camera coordinate system to the laser coordinate system. Section III gives methods for solving for the extrinsic calibration, first showing a direct solution, followed by nonlinear iterative optimization methods which use the results from previous steps as initial conditions. A global optimization is also proposed to refine both camera intrinsic and extrinsic parameters. We conclude by giving experimental results showing the success of the techniques presented.

II. BASIC EQUATIONS

A camera can be described by the usual pinhole model. A projection from the world coordinates $P = [X, Y, Z]^\top$ to the image coordinates $\mathrm{p} = [u, v]^\top$ can be represented as follows [6]:

$$\mathrm{p} \sim \mathrm{K}(RP + \mathrm{t}) \tag{1}$$

where K is the camera intrinsic matrix, R a 3×3 orthonormal matrix representing the camera's orientation, and t a 3-vector representing its position. In real cases, the camera can exhibit significant lens distortion, which can be modelled as a 5-vector parameter consisting of radial and tangent distortion coefficients. We assume in the remainder of this paper that the camera has no significant lens distortion, or that the images have already been warped to eliminate it.

The laser range finder reports laser readings which are distance measurements to the points on a plane parallel to the floor. A laser coordinate system is defined with an origin at the laser range finder, and the laser scan plane is the plane $Y = 0$. Suppose a point P in the camera coordinate system is located at a point P^f in the laser coordinate system, and the rigid transformation from the camera coordinate system to laser coordinate system can be described by :

$$P^f = \Phi P + \Delta \tag{2}$$

where Φ is a 3x3 orthonormal matrix representing the camera's orientation relative to the laser ranger finder and Δ is an 3-vector corresponding to its relative position.

Our goal in this paper is to develop ways to solve for these extrinsic camera parameters Φ and Δ which define the position and orientation of the camera with respect to the laser coordinate system.

A. Geometric Constraints

Our proposed calibration method is to place in front of our system a planar pattern, say a checkerboard, which is visible to both the camera and the laser ranger finder. Figure 1 provides a general setup of this calibration method. For simplicity, when we talk about a *calibration plane*, we refer to the plane surface defined by the checkerboard. And we use *laser points* to refer to the laser measurements on the checkerboard, which is a portion of the whole laser reading.

Without loss of generality, we assume that the calibration plane is the plane $Z = 0$ in the world coordinate system. In the camera coordinate system, the calibration plane can be parameterized by 3-vector N such that N is parallel to the normal of the calibration plane, and its magnitude, $\|N\|$, equals the distance from camera to the calibration plane. Using (1) we can derive that

$$N = -R_3(R_3^\top \cdot \mathrm{t}) \tag{3}$$

where R_3 is the 3rd column of rotation matrix R, and t the center of the camera, in world coordinates.

Since the laser points must lie on the calibration plane estimated from the camera, we get a geometric constraint on the rigid transformation between the camera coordinate system and the laser coordinate system. Given a laser point P^f in the laser coordinate system, from (2), we can determine its coordinate P in the camera reference frame as $P = \Phi^{-1}(P^f - \Delta)$. Since the point P is on the calibration plane defined by N, it satisfies that $N \cdot P = \|N\|^2$. Then we have

$$N \cdot \Phi^{-1}(P^f - \Delta) = \|N\|^2 \tag{4}$$

For a measured calibration plane parameters N and laser point P^f, this gives a constraint on Φ and Δ.

III. SOLVING EXTRINSIC CALIBRATION

This section provides the details how to effectively solve the extrinsic calibration problem for the system of a camera and laser range finder. We will first propose a linear solution, followed by a nonlinear optimization with outlier detection. Finally, a global optimization is performed to refine the extrinsic camera parameters. This can be extended to refine the intrinsic parameters more accurately than the standard singe camera calibration method.

A. Linear Solution

First, we assume the camera is calibrated [2] and what remains is to determine the calibration plane parameters by solving the pose of the camera with respect to the checkerboard, which is discussed in [13]. Once the camera's extrinsic parameters (R, t) are determined with respect to the checkerboard, the calibration plane parameter N can be obtained by (3).

Since all laser points are on the plane Y=0 in the laser coordinate system, a laser point P^f can be represented by $\hat{P}^f = [X, Z, 1]^\top$. Then we rewrite Equation (4) as:

$$N \cdot H \hat{P}^f = \|N\|^2 \tag{5}$$

where $H = \Phi^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 0 & -\Delta \\ 0 & 1 \end{pmatrix}$, a 3x3 transform matrix from the laser coordinate system to the camera coordinate system. For each pose of the calibration plane. we have several linear equations in the unknown parameters of H, which we solve with standard linear least squares.

Once H is determined, we can estimate camera relative orientation and position as follows:

$$\Phi = [H_1, -H_1 \times H_2, H_2]^\top$$
$$\Delta = -[H_1, -H_1 \times H_2, H_2]^\top H_3$$

where H_i is the ith column of matrix H.

This computed matrix Φ may not meet the properties of a rotation matrix. A rotation matrix $\hat{\Phi}$ can be calculated to approximate the computed matrix Φ by minimizing Frobenius norm of the difference $\hat{\Phi} - \Phi$, subject to $\hat{\Phi}\hat{\Phi}^\top = I$. The details about this matrix computation can be referred in [4]. This concludes the direct estimation of $\hat{\Phi}$ and Δ, the relative pose of the camera with respect to the laser range finder.

B. Nonlinear Optimization

The above solution is obtained by minimizing an algebraic distance which is not directly related to our measurements. We can refine it by a nonlinear minimization on the Euclidean distances from laser points to the checkerboard planes, which is more physically meaningful.

Equation (4) gives the Euclidean distance between a laser point and the calibration plane. Given different poses of the checkerboard, we define an error function $f(\Phi, \Delta)$ as the sum of these distances for every laser point j in very data set i:

$$\sum_i \sum_j \left(\frac{N_i}{\|N_i\|} \cdot (\Phi^{-1}(P_{ij}^f - \Delta)) - \|N_i\| \right)^2 \quad (6)$$

where N_i defines the checkerboard plane in the ith pose. A rotation Φ is parameterized by the Rodrigues formula as a 3-vector parameter, which is in the direction the rotation axis and has a magnitude equal to the rotation angle. We minimize (6) as a nonlinear optimization problem by using the Levenberg-Marquardt method [8], [10], [11]. This requires an initial guess of Φ and Δ, which is obtained using the method described in the previous section.

Both camera and laser range finder have some noises in their outputs, and we found that the noise in laser points and wrong estimates of calibration planes affect the final results. We develop a robust method that iteratively throws away some data sets which have too large residuals from (6). A brief description is given as follows:

1) For each checkerboard view, the calibration plane parameter N is estimated in the camera coordinate system, and the laser points are extracted in the laser coordinate system. The noise in the extracted laser points is also roughly estimated by fitting line to these points. Construct a valid pose set with all checkerboard views.

2) Based on current valid pose set, estimate camera orientation Φ and position Δ by minimizing the weighted version of (6), which can be written as:

$$\sum_i \sum_j w_i \left(\frac{N_i}{\|N_i\|} \cdot (\Phi^{-1}(P_{ij}^f - \Delta)) - \|N_i\| \right)^2 \quad (7)$$

where w_i is a weight for the laser points in the ith pose, and it can be calculated based on the estimated noise in the laser points.

3) Empty the valid pose set. For each pose, compute the average 3D projection error ϵ of laser points to the checkerboard plane with current estimated camera orientation and position. If the projection error $\epsilon < \delta$, then we add this pose into the valid pose set, where δ is threshold for maximum average error.

4) Repeat Step 2, 3 until convergence (empirically it takes 2 to 3 iterations to converge).

C. Global Optimization

In practice, the camera calibration is not precisely known, and measurement errors affect the performance of the extrinsic calibration. Given the relative orientation and position of a camera, the laser data gives extra constraints on the position of the planar pattern, which could be analyzed in the camera intrinsic calibration. So intuitively, the camera intrinsic and extrinsic parameters can be refined by performing a global optimization with the initial estimate of the extrinsic camera parameters Φ and Δ, and an estimate of the intrinsic parameter K.

Given different views of the checkerboard and grid points on the plane, We can obtain the projection error of grid points:

$$\sum_i \sum_j \|\mathbf{p}_{ij} - \tilde{\mathbf{p}}(K, R_i, t_i, P_j)\| \quad (8)$$

where $\tilde{\mathbf{p}}(K, R_i, t_i, P_j)$ is the projection of point P_j in image i according to (1).

With (6), we can do a global optimization by minimizing the combination of reprojection error and laser to calibration plane error as follows:

$$\sum_i \sum_j \mathrm{d}^2(P(\Phi, \Delta, P_{ij}^f), N(R_i, t_i)) +$$
$$\alpha \sum_i \sum_j \|\mathbf{p}_{ij} - \tilde{\mathbf{p}}(K, R_i, t_i, P_j)\|^2 \quad (9)$$

where $N(R_i, t_i)$ the calibration plane parameter in pose i according to (3), $P(\Phi, \Delta, P_{ij}^f)$ the coordinate of laser point P_{ij}^f in the camera coordinate system according to (2), $\mathrm{d}^2(P, N)$ the squared Euclidean distance from point P to plane N, and α is a scalar weight to normalize the relative contribution of the laser error function and the camera error function.

This nonlinear minimization problem can be solved with the Levenberg-Marquardt method. It demands the initial guesses of Φ, Δ, and $\{R_i, t_i \mid i = 1 \ldots n\}$, which can be obtained with the methods discussed in the previous sections, and an initial guess of camera intrinsic matrix K.

D. Algorithm Summary

The complete algorithm can be described as the following steps:

1) Build a big checkerboard and place it in front of the camera-laser range finder system in the different orientations.
2) For each checkerboard pose, extract the laser points in the laser reading, and detect the checkerboard grid points in the image. Estimate the camera orientation R_i and position t_i with respect to the checkerboard, and then compute the calibration plane parameter N_i.
3) Estimate the parameter Φ and Δ using the linear solution as described in Section III-A.
4) Refine Φ and Δ using the techniques described in Section III-B.
5) If necessary, refine all parameters by minimizing (9).

IV. EXPERIMENTS

The proposed algorithm has been implemented in Matlab and is available on request to the author. This section illustrates tests on both computer simulated data and real data. The closed-form solution provides initial conditions for the nonlinear refining. A global optimization on all camera calibration parameters is also tested.

A. Computer Simulations

The placement of the camera relative to the laser range finder is described by the camera's position $\Delta = [\ 0,\ -1.0,\ 0.1\ ]^\top$ meters and orientation Φ defined by a rotation vector $[\ -0.25,\ -0.02,\ -0.01\ ]^\top$.

The camera is simulated as an ideal pinhole with focal length 750 and principal point $(320, 240)$. The calibration pattern plane is a checkerboard defined by 10×10 grids, and the size of the pattern square is $76mm \times 76mm$. The orientation of the checkerboard is generated as follows: the plane is initially parallel to the image plane; a rotation axis is randomly chosen on the plane and the plane is rotated around that axis with angle θ. The position of the plane is chosen properly such that the checkerboard grids can appear entirely on the image plane. Gaussian noise with mean 0 and standard deviation 0.5 pixel is added to the projected image points. The laser points are computed based on the placement of the camera and the setting of checkerboard. We also add uniform noise into the laser points of $\pm 5cm$ which is approximately the same as the observed noise distribution in our sensor.

In the experiment, the estimated extrinsic parameters are compared with the ground truth. We measure the error for camera orientation Φ by computing the angle between the estimate and the true orientation, and the error for camera position Δ by computing the distance between the estimate and the true camera position.

Performance w.r.t. the number of checkerboard poses. This experiment deals with how the number of plane poses effects the performance. We vary the number of poses from 6 to 24. For each experiment number, 100 trials of independent checkerboard plane orientations with $\theta = 60°$ and independent checkerboard plane positions are conducted. The Gaussian noise added in the projected image points is also independent between trials, as well as the uniform noise in the laser points. The results is shown
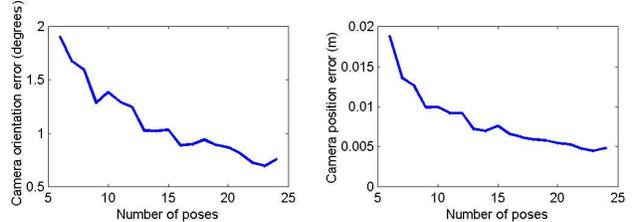

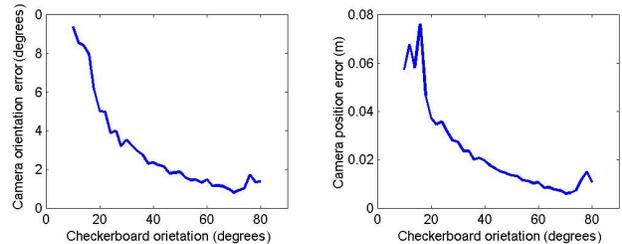Fig. 2. Errors vs. the number of poses of the checkerboard plane

Fig. 3. Errors vs. the orientation of the checkerboard plane

in Figure 2. The error decreases when more poses are used.

Performance w.r.t. the orientation of checkerboard plane. This experiment is performed for different orientations of the checkerboard plane to examine its influence in the calibration performance. We compute the average errors by running 100 trials with 10 checkerboard poses. The orientation angle of checkerboard plane varies from $10°$ to $80°$, and the result is shown in Figure 3. We found that the result improves when the orientation angle increases, due to more precise estimates of calibration planes with large angles with respective to the image plane. Best performance seems to be achieved around $70°$. Note that in practice, when the angle increases, the number of laser points probably decreases, and foreshortening makes the corner detection less accurate, but these are not considered in this experiment.

Refinement of camera intrinsic parameters. This experiment investigates how the laser data help the camera calibration, and examines its performance on refining camera intrinsic and extrinsic parameters. Here we run 100 independent trials and then compare the average result with the ground truth. For each trial, 10 checkerboard poses are used. To randomize starting conditions, the camera focal length is corrupted with Gaussian noise with mean 0 and standard deviation 10 pixels, and the camera principal point is corrupted with Gaussian noise with mean 0 and deviation 5 pixels. The checkerboard is placed in the orientation ranging from $50°$ to $70°$. Initially, we perform the extrinsic calibration with the corrupted camera internal matrix K, which gives incorrect estimates of the relative position Δ and orientation Φ. Table I shows the accuracy of these Φ, Δ estimates before and after global optimization, showing that we recover somewhat from corrupted camera internal matrix K. The improvement of camera intrinsic matrix can be evaluated by the ratio of the Frobenius norm of the
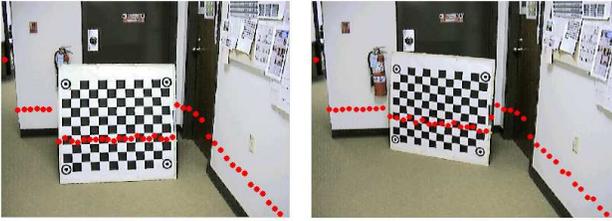
Fig. 4. The figure shows 2 out of 10 checkerboard settings captured by the camera. The laser range finder points are projected on the image using the computed Φ and Δ.

difference of the estimated K and the ground truth to the Frobenius norm of the difference of the corrupted K and the ground truth. As we can see, when the global optimization is applied on both camera intrinsic and extrinsic parameters, the accuracy of the camera intrinsic matrix K is improved by about 30%. The error of estimates of camera orientation and position with respect to the laser coordinate system also decreases by around 30%.

	initial	final
Orientation error(Φ)	2.33°	1.95°
Position error(Δ)	3.78cm	2.37cm
Fro. norm ratio w.r.t K	0.6969	

TABLE I

THE RESULT OF GLOBAL OPTIMIZATION ON THE CAMERA INTRINSIC AND EXTRINSIC PARAMETERS

B. Real Data

The proposed method has been tested on a robotic platform illustrated in Figure 1, equipped with a SICK-PLS laser range finder and a Sony DFW-VL500 digital camera mounted on top of the robot. The laser range finder provides range measurements by scanning 180 degrees of the environment parallel to the floor, with an angle resolution of one measurement per degree and a range measuring accuracy of ± 5cm. On real data, the method operates well given reliable calibration parameters of the camera. Here we present the result with one example.

The camera resolution is set as 640×480. The calibration pattern is a 12×10 checkerboard, and the size of a checker square is $76mm \times 76mm$. 10 images of the checkerboard are taken along with 10 laser readings. The laser points can be manually selected among the whole laser measurements. Figure 4 demonstrates the results of the algorithm applied to the configuration. We map the laser points onto the calibration plane with estimated Φ and Δ, and the average distance error from the laser points to the calibration plane is around 2-3cm. We do not have the ground truth of the extrinsic parameters Φ and Δ, but the mapping results are quite reasonable.

C. Refinement of intrinsic camera parameters

In addition, our global optimization refines the intrinsic camera parameters, and we consider this optimization for

the camera calibration of this same data set as Section IV-B. An initial solution for intrinsic camera parameters is made from the image data alone. This is listed as the "initial" column in Table II, which illustrates the estimation of the focal length (f_x,f_y) and the principal point (c_x,c_y). After solving for Φ, Δ and then applying global optimization, the intrinsic camera parameters change and are shown in the "final" column. The variance of these estimates is calculated by a leave-one-out method. Given the 10 data samples, the estimator is run 10 times, each time ignoring a different data sample. The variance of these 10 runs is reported as σ. While the estimated values remain very similar, there is a small, but consistent, improvement in the variance of the estimates.

	initial	σ	final	σ
f_x	768.60	4.34	768.51	3.88
f_y	768.11	4.55	768.05	4.04
c_x	319.90	6.74	319.27	6.65
c_x	268.49	9.35	268.88	9.19

TABLE II

THE RESULT OF GLOBAL OPTIMIZATION ON THE INTRINSIC CAMERA PARAMETERS

V. CONCLUSION

In this paper, we presented an extrinsic calibration method to estimate the orientation and position of a camera with respect to a 2D laser ranger finder for the robot. The proposed method requires a few poses of planar pattern which is visible for both the camera and the laser range finder, and then a geometric constraint on the extrinsic camera parameters is imposed. This calibration succeeds and applies to a very common sensor package for a large number of autonomous robots, such as the iRobot series. Moreover, the camera intrinsic calibration can be also improved with laser data, which we believed is helpful for many robotic vision tasks.

REFERENCES

[1] H. Baltzakis, A. Argyros, and P. Trahanias. Fusion of laser and visual data for robot motion planning and collision avoidance. *Machine Vision and Application*, 12:431–441, 2003.
[2] Jean-Yves Bouguet. Camera Calibration Toolbox for Matlab, 2003.
[3] C. Frh and A. Zakhor. Data processing algorithms for generating textured 3d building facade meshes from laser scans and camera images. In *Proc. 3D Data Processing, Visualization and Transmission 2002*, pages 834 – 847, June 2002.
[4] G. Golub and C. Van Loan. *Matrix Computation*. John Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, Maryland, third edition, 1996.
[5] C. Grimm and B. Smart. Lewis the Robot Photographer. In *ACM SIGGRAPH*, San Antonio, Texas, 2002.
[6] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2000.
[7] O. Jokinen. Self-calibration of a light striping system by matching multi-ple 3-d profile maps. In *Proc. the 2nd International Conference on 3D Digital Imaging and Modeling*, pages 180–190, 1999.
[8] K. Levenberg. A method for the solution of certain problems in least squares. *Quarterly Applied Math. 2, pp. 164-168, 1944.*, 2:164–168, 1944.
[9] Y. Liu and R. Emery. Using EM to learn 3d environment models with mobile robots. In *Proc. 18th International Conference on Machine Learning*, 2001.

[10] D. Marquardt. An algorithm for least squares estimation of nonlinear parameters. *SIAM Journal of Applied Math*, 11:431–441, 1963.

[11] J. J. Mor. The Levenberg-Marquardt algorithm: Implementation and theory. In G A Watson, editor, *Lecture Notes in Mathematics*, volume 630, pages 105–116. Springer Verlag, 1977.

[12] D. Ortin, J. Montiel, and A. Zisserman. Automated multisensor polyhedral model acquisition. In *Proc. IEEE International Conference on Robotics and Automation*, pages 1007–1012, 2003.

[13] Zhengyou Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *Proc. International Conference on Computer Vision*, September 1999.