Fully Automated Laser Range Calibration

Matthew Antone and Yuli Friedman
Computer Vision Group
BAE Systems Advanced Information Technologies
Burlington, MA, USA
{matthew.antone,yuli.friedman}@baesystems.com

Abstract

We present a novel method for fully automated exterior calibration of a 2D scanning laser range sensor that attains accurate pose with respect to a fixed 3D reference frame. This task is crucial for applications that attempt to recover self-consistent 3D environment maps and produce accurately registered or fused sensor data.

A key contribution of our approach lies in the design of a class of calibration target objects whose pose can be reliably recognized from a single observation (i.e. from one 2D range data stripe). Unlike other techniques, we do not require simultaneous camera views or motion of the sensor, making our approach simple, flexible and environment-independent.

In this paper we illustrate the target geometry and derive the relationship between a single 2D range scan and the 3D sensor pose. We describe an algorithm for closed-form solution of the 6 DOF pose that minimizes an algebraic error metric, and an iterative refinement scheme that subsequently minimizes geometric error. Finally, we report performance and stability of our technique on synthetic and real data sets, and demonstrate accuracy within 1 degree of orientation and 3 cm of position in a realistic configuration.

1 Introduction

In recent years, the ubiquity of laser range sensors (*lidars*) has increased, and their application to many domains – including vision and robotics – has grown rapidly. Indeed, heterogeneous sensor suites consisting of multiple lidars, cameras, and other devices have become quite common for such applications as object recognition, tracking, navigation, and environment reconstruction.

For multi-sensor systems to be useful, they must produce measurements in a common coordinate system so that observations from two or more sensors may be related to one another in a meaningful way within a consistent *relative* reference frame. Further, for higher-level geometric reasoning it is often essential for sensors to be situated with respect to a known *absolute* reference frame. In simultaneous localization and mapping, for instance, sensors must be carefully calibrated with respect to the robot's body frame to enable the robot to accurately localize obstacles and landmarks (Figure 1).

Several different types of lidar devices currently exist, with varying scanning mechanisms, number of lasers, and geometric configurations. In this work we focus on the most
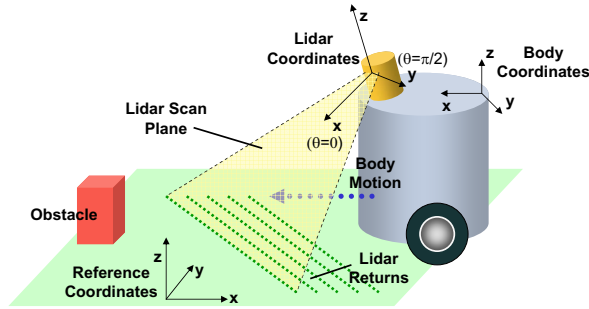
Figure 1: A typical application in which a lidar is rigidly attached to a mobile robot that produces pushbroom-like scans of its environment. We wish to determine the transformation between the lidar's coordinate system and that of a known reference such as the body frame or the environment.

common, the planar or line-scanning lidar. The calibration of absolute orientation is especially important and challenging for these sensors, as they produce only 2D slices of their 3D environment. As the lidar plane moves (e.g. spins at a fixed location to produce full spherical scans, or translates on a mobile robot to sweep out the environment ahead), systems require accurate hand-eye calibration to assemble the individual slices into a metric – and thus physically meaningful – 3D point cloud.

In this paper we present a novel and fully-automated procedure for calibrating the exterior 6-degree-of-freedom pose, consisting of 3D location and orientation, for a planar lidar sensor with respect to a fixed reference frame. We make no assumptions about the surrounding environment and do not require the sensor to physically move. Our algorithm merely requires a single range-only scan of a carefully-designed calibration target, making the technique flexible and independent of errors in other sensor measurements such as odometers or servos.

1.1 Prior Work

There are many strategies for calibrating the pose of a lidar sensor. In the literature, these tend to fall into two main categories: those that require the use of additional sensors such as cameras, and those that require known motion of the lidar itself. Multi-sensor calibration typically determines the relative Euclidean transformation between the lidar and a rigidly-attached camera. Zhang and Pless present a technique in which both sensors image a planar checkerboard target at different (unknown) orientations; the camera's pose with respect to the target is determined using a standard extrinsic calibration method, and combined with straight-line profiles extracted from the lidar to form constraints on the lidar's pose [10]. Mei and Rives developed a more general theory of lidar registration to catadioptric cameras, considering additional cases in which the laser returns are visible to the camera, thus forming explicit correspondences between image pixels and range samples [6]. Several methods have also been proposed that utilize structured light, or visible laser profiles in the images [4, 7].

Motion-based or *active* calibration involves imaging objects from multiple locations and orientations. Several methods move a robotic arm, to which the laser is attached, and

capture range data of a fixed planar surface at different orientations; knowledge of precise relative arm pose is assumed [8, 2]. McIvor places a cubic calibration target on a motion table, and "scans" the object with the lidar to obtain both range and intensity data; the laser pose that best rectifies the data to the known target geometry is then determined [5]. Finally, Zhang and Pless estimate egomotion of a moving robot via robust matching of consecutive lidar scans over time, using the structure of the surrounding environment to constrain the hand-eye lidar pose [9].

1.2 Contributions

In this work, we develop a flexible lidar calibration technique based on a novel target object design. Our technique has several unique characteristics that provide advantages over prior methods: it does not require additional sensors such as cameras; it does not require motion of the sensor or the target; and it requires only range data, not intensity data. These properties offer applicability to a wide variety of different 2D laser scanners and imaging geometries. Further, because our target's surface can be painted with visible patterns, our method also enables precise lidar-camera registration. To our knowledge, this is the first such published technique.

2 Fundamentals

We seek the rigid-body Euclidean transformation (R, T) that aligns a particular sensor's local coordinate system with some fixed reference frame. Here, T is the sensor's 3D location in the reference frame, and the columns of the orthonormal rotation matrix R represent the axes of the sensor expressed in reference coordinates. Thus, the transformation between a reference point M and a sensor point Q is given by

$$Q = R^T (M - T),\qquad(1)$$

and the inverse transformation is given by

$$M = RQ + T.\qquad(2)$$

The laser produces 2D range scans consisting of discrete angular samples in a plane. For each ray sample along the direction denoted by angle θ_k, the sensor measures a corresponding distance r_k along that ray. We place the origin of the sensor coordinate system at the origin of all laser rays; the ray at $\theta = 0$ is coincident with the sensor's x axis, and the ray at $\theta = \pi/2$ is coincident with the y axis. Thus, the scan produces measurements of the form

$$Q_k^T = \begin{pmatrix} u_k & v_k & 0 \end{pmatrix} = r_k \begin{pmatrix} \cos\theta_k & \sin\theta_k & 0 \end{pmatrix}.\qquad(3)$$

Since the range measurements all lie in a plane, the transformation between lidar coordinates and reference coordinates given in (2) can be reduced to a homography [3] as

$$M = Hq = RKq = R \begin{pmatrix} e_x & e_y & R^T T \end{pmatrix} q,\qquad(4)$$

where e_x and e_y represent unit vectors in the x and y direction, respectively, and $q^T = \begin{pmatrix} u & v & 1 \end{pmatrix}$.

Internal mis-calibration and servo scan inconsistency may cause error in θ, which is generally negligible. Range measurements r also exhibit uncertainty induced by quantization (i.e. fixed bit allocation in analog-to-digital conversion of the range return), typically on the order of 1cm, and by measurement noise from material reflectivity and environmental effects, typically also on the order of 1cm [1]. Finally, nonlinear range errors can arise from depth discontinuities in imaged surfaces due to non-zero laser spot size.

3   Calibration Target Design

The most important factor in designing a target object suitable for single-scan range-based calibration is the 3D pose recovery itself. While many simple objects allow determination of certain straightforward extrinsic parameters, the design problem becomes more challenging as the number of required DOFs increases. A second factor is reliable and stable detection of the target object, which must account for lidar measurement errors and discrete rather than continuous angular sampling. Finally, there are practical matters to consider such as the ease, repeatability, and accuracy of physical target construction.

The lidar produces as measurements a series of ranges at specified angles that constitute the (sampled) intersection of a virtual plane with the visible 3D surfaces in the scene. We next consider two classes of simple 3D objects that produce unique 2D cross sections determined entirely by the orientation and position of the slicing plane.

3.1   Conic Sections

It is well known that conic sections are produced by intersections of a plane with a double cone, with the particular type and shape of the curve defined entirely by the position and orientation of the slicing plane. For simplicity, and without loss of generality we assume a cone with apex at the origin and whose axis of symmetry is the z axis in reference coordinates, defined by the implicit equation

$$X^2 + Y^2 - Z^2 = 0 = M^T S M, \tag{5}$$

where the reflection matrix $S = diag(1, 1, -1)$. Intuitively, the size of a cross section is related to the translation of the sensor slice plane along the symmetry axis, and the shape of the cross section is related to the slice plane's orientation (Figure 2).

Substitution of (4) into (5) reveals

$$(Hq)^T S(Hq) = 0 = q^T K^T R^T S R K q. \tag{6}$$

Thus, each observed lidar sample generates a single quadratic constraint equation in the 6 unknown parameters encoded by R and T, and it would seem that six such points in non-degenerate configuration uniquely determine these parameters. However, due to inherent geometric symmetries, not all DOFs may in fact be recovered.

A geometric illustration of this ambiguity is shown in Figure 2. Algebraically, we can see that applying a rotation about z to reference point M does not affect (6) since the reflection and rotation axes are identical; we can nonetheless obtain an accurate estimate for R and T in closed form, modulo this rotation about z (see Appendix A). Unfortunately, this inherent ambiguity precludes use of the cone for full 6 DOF calibration. Further, a conic target is difficult to construct with sufficient size and precision for pose recovery.
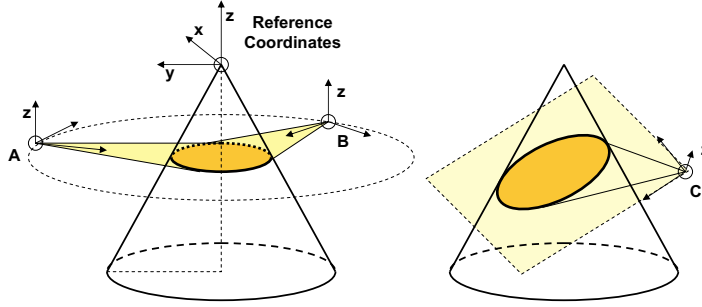
Figure 2: A conic section orthogonal to the symmetry axis forms a circle whose size varies with z. Lidars with poses A and B would produce the same circular scan (left). An off-axis slice plane produces elliptical or hyperbolic cross sections determined by the plane's orientation (right).
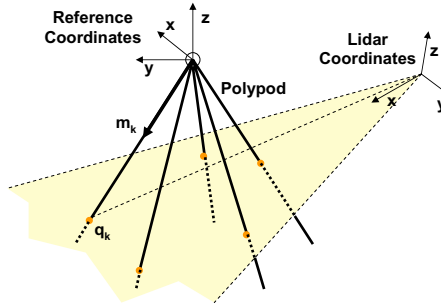

Figure 3: Polypod legs with direction m_k intersect the lidar slice plane to form points q_k in sensor coordinates. The relationship between m_k and q_k is a plane projective homography.

3.2 Polypods

We define a *polypod* as a multi-legged structure consisting of a set of rays that emanate from a single point in known directions. Without loss of generality we assume that the ray origin is coincident with the origin of the reference coordinate system. As with the cone, moving the slice plane along the z axis changes the size of the cross section, and tilting the plane changes the cross section's shape (Figure 3). However, unlike the cone, a polypod can be designed so that its cross section uniquely determines the sensor's full 6 DOF pose, up to a duality of solutions.

We now derive a geometric relationship between the polypod legs and the lidar measurements given a particular orientation R and position T. Let $m_k^T = \begin{pmatrix} x_k & y_k & 1 \end{pmatrix}$ represent the direction of the kth ray, or polypod leg; reference points M on the 3D ray may thus be parameterized as $M = m_k s$. To determine the lidar points Q, we then transform the ray to lidar coordinates and find its intersection with the lidar slice plane $z = 0$. The new ray direction is given by $R^T m_k$, and the new ray origin is $-R^T T$, according to the transformation (1). Thus, the ray parameterization becomes

$$Q = -R^T T + R^T m_k s. \tag{7}$$

To find the intersection of the ray and the plane $z = 0$, we determine the parameter s such that the z component of Q is zero. Solving (7) for $Q^T e_z = 0$ gives $s = R_z^T T / R_z^T m_k$, where R_z is the third column of the rotation matrix R. Substituting back into (7) gives

$$Q = R^T (m_k \frac{R_z^T T}{R_z^T m_k} - T), \qquad (8)$$

which we can re-write component by component as

$$\begin{pmatrix} u_k \\ v_k \\ 1 \end{pmatrix} R_z^T m_k = q_k R_z^T m_k = \begin{pmatrix} R_x^T m_k R_z^T T - R_x^T T R_z^T m_k \\ R_y^T m_k R_z^T T - R_y^T T R_z^T m_k \\ R_z^T m_k \end{pmatrix}. \qquad (9)$$

After manipulation of (9), we find that

$$q_k \sim \begin{pmatrix} t_z & 0 & -t_x \\ 0 & t_z & -t_y \\ 0 & 0 & 1 \end{pmatrix} R^T m_k = K^{-1} R^T m_k \qquad (10)$$

where $t = R^T T$. The 3×3 matrix $K^{-1} R^T$ defines a plane projective homography between measured points q_k on the lidar slice plane and the corresponding polypod rays m_k.

Note that (10) is the inverse of (4), the basic relationship between lidar points and reference points; here, polypod legs equate to projective rays. Further, we observe that the ambiguity from symmetry in the conic case is completely resolved, because cross sections are not rotationally symmetric and because there is a unique homography that relates points to points, rather than points to surfaces, provided that at least four non-degenerate correspondences exist [3].

3.3 Pyramid Target

Having derived the abstract geometric relationship between polypod configuration and lidar cross section, we now define a more practical and concrete calibration target design. First, for precise and simple construction, and because a homography is uniquely defined by four correspondences, we use the minimum number of four legs. We also avoid construction of a literal polypod; noisy measurements of its thin legs would lead to significant errors in recovered pose, and in fact the legs might be missed by the laser entirely.

Rather than detecting the polypod legs directly as single-point measurements, then, we design our target so that we can indirectly, but more reliably, infer their locations in the scan. To achieve this, we construct a pyramid-shaped target whose four planar faces intersect to form "virtual" polypod legs (Figure 4). The cross section is a quadrilateral whose edges may be estimated directly from sets of multiple scan points, and whose vertices q_k may then be recovered more precisely as the intersections of these edges.

4 Pose Recovery

To reduce the effects of quantization error and stochastic noise in range and angle measurements, we keep the target stationary in the lidar's field of view for a few seconds and
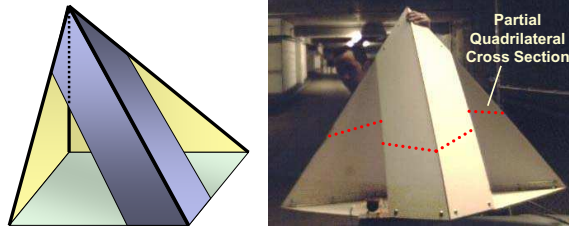
Figure 4: Our target design consists of a pyramid (left) with holes cut into the front faces so that the back faces are visible. Bold lines indicate "virtual" polypod legs formed by intersection of the pyramid faces. A lidar scan of this object (right) forms a quadrilateral cross section that uniquely determines the sensor pose.
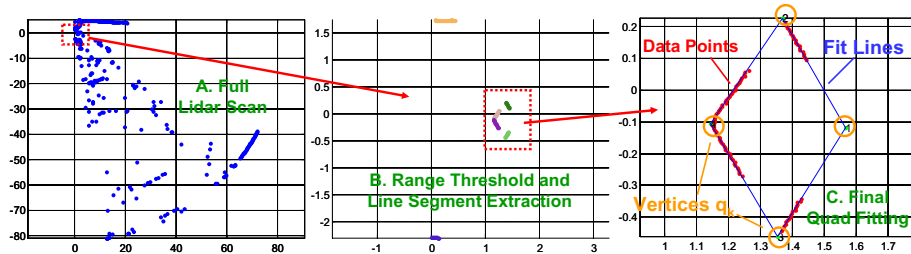

Figure 5: Target segmentation involves several steps that result in rejection of clutter points and a robustly estimated quadrilateral cross section of the target object.

obtain a series of several hundred scans, which are then averaged to form a single, noise-suppressed scan. Our algorithm processes this data to automatically segment the target points from the background, find the vertices of the quadrilateral cross section, solve for the approximate pose in closed form, and iteratively refine the pose for higher accuracy.

4.1 Segmentation

A particular lidar scan contains both the desired cross section of the calibration target object and clutter from the environment. Our first task is therefore to segment the points of interest from the clutter (Figure 5). We first remove all points beyond a threshold distance from the sensor. We next search the scan for contiguous straight line segments using a successively applied RANSAC algorithm [3]. Each RANSAC iteration selects two points at random from the scan, fits a line to those points, and evaluates the remaining points against the line. The line associated with the most inliers is kept, the inliers are removed from the scan points, and the procedure repeats until no additional lines are found.

We next subdivide each line into contiguous segments, removing segments shorter than a threshold length, where the threshold is related to the target's expected cross sectional dimensions. The final step is to find the "correct" four line segments corresponding to the true target cross section. We first search for candidate segment pairs that might form the front two faces; we then search the remaining segment pairs for closed quadrilaterals.

4.2 Position and Orientation Estimation

We form the four vertices q_k via intersection of the extracted line segments, then apply an ordering constraint to assign the q_k to the correct (known) polypod leg directions m_k. Having established four correspondences, we solve for the homography $H = RK$ using a standard technique [3].

We next factor H using QR decomposition, which results in an orthonormal matrix R and an upper-triangular $\tilde{K}$. Because of the unknown scale on H, the matrix $\tilde{K}$ will not be of the proper form of (4). In the absence of noise, it will differ by a constant scale factor, so we divide $\tilde{K}$ by either of its first two diagonal entries (which are equal) and determine T by left-multiplying the third column of the result by $-R$. In general, however, the q_k will be noisy, and $\tilde{K}$ will not simply differ from K by a constant scale factor. We therefore approximate the true K by dividing $\tilde{K}$ by the average of its first two diagonal entries (which are no longer equal), and determining T as before. Note that this differs from the factorization of Zhang [11], which solves for K exactly and then approximates R by finding the "closest" rotation in Frobenius norm to the estimated $\tilde{R}$.

In general, when noise is present in the data, the above algebraic approximation leads to a pose solution that is reasonably accurate, but not optimal. This solution can, however, be used to initialize a direct optimization method on R and T, such as Levenberg-Marquardt, that seeks to minimize the sum of Euclidean distances between the transformed points Hq_k and their counterparts m_k.

5 Experiments

To demonstrate the efficacy of our target object design and calibration method, we conducted several different experiments in simulation (to systematically control noise sources and other DOFs) and on real data (to show performance with physical sensor, target, and measurements). Results of these experiments are presented below.

5.1 Synthetic Data

We generated simulated data for experiments as follows. First, we constructed a virtual 3D model of the target pyramid with realistic dimensions (approximately one meter on a side). Next, we chose a set of lidar poses with translations varying between 2 and 30 meters and rotations varying between 0 and 45 degrees about each axis. For each pose we generated virtual lidar scans by intersecting scan rays with the target surface at 0.25 degree spacing. Ray angles and intersection ranges were perturbed by additive Gaussian noise with specified variance; ranges were also quantized to 1cm levels, and range discontinuities were averaged to simulate real sensor phenomena mentioned in Section 2. Results are shown in Figure 6; we report the mean position and orientation errors over 500 trials for each variant.

5.2 Real Data

We constructed a physical calibration target (Figure 4) and configured six different sensor poses with the target sitting on the ground. We measured the lidar's position and orientation by hand, so the "ground truth" pose was known only approximately. The sensor was
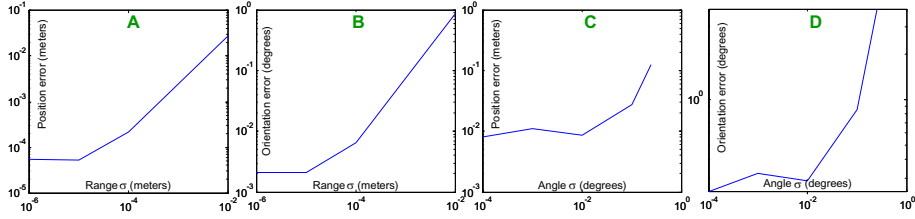
Figure 6: Performance of our algorithm on simulated data, plotting position and orientation error with varying range uncertainty (A and B) and angular uncertainty (C and D).

Table 1: Position and Rotation Error using Real Data

Dist (cm)	Roll (°)	Pitch (°)	Yaw (°)	Pos Err (cm)	Rot Err (°)
356 ± 5	0 ± 2	0 ± 2	30 ± 2	3.1	1.1
311 ± 5	0 ± 2	30 ± 2	15 ± 2	5.4	1.8
347 ± 5	10 ± 2	0 ± 2	45 ± 2	4.9	0.9
312 ± 5	10 ± 2	30 ± 2	45 ± 2	6.1	1.3
623 ± 5	10 ± 2	30 ± 2	45 ± 2	2.2	0.5
988 ± 5	10 ± 2	30 ± 2	45 ± 2	3.2	0.4

a SICK LMS291 set for $0.25°$ angular and 1cm range resolution. At each pose, we collected and averaged several hundred scans of the object, then ran our algorithm to segment the target and solve for R and T. Results are summarized in Table 1.

6 Conclusions and Future Work

We have presented a unique target object design and algorithm for automatic calibration of a 2D laser range sensor. Unlike previous methods, ours requires neither additional sensors such as cameras nor motion of the lidar, affording a great deal of flexibility and generality. We have derived target geometries that support estimation of extrinsic pose from a single cross sectional range measurement, and suggested a specific pyramid design based on a quadropod. Our techniques were demonstrated to exhibit robustness and accuracy, reliably locating the target amidst clutter and estimating pose to within less than $1°$ of rotation error and a few cm of position error for realistic sensor characteristics.

Our technique relies on precise construction of the calibration target, and requires that the scan's slice plane fall within a valid band that intersects all four faces of the pyramid. Because of finite angular resolution, the target must be placed within a small enough radius (empirically on the order of 20 meters) that it produces a sufficient number of samples on the target surface for reliable estimation. In practice, the target also must be held still so that range errors may be diminished via averaging of multiple scans.

We are currently working to incorporate this technique into a larger end-to-end system for self-consistent calibration of heterogeneous multi-modal sensor suites. For true absolute pose estimation – e.g. relative to an inertial frame rather than to the calibration target – a precise and repeatable method must be developed to measure the target's pose with respect to the frame of interest. Our pyramid target could be modified to simultaneously calibrate cameras by placing visible fiducials or patterns on its surface. Finally, it would be interesting to study the effect of relative target dimensions on pose estimation.

References

[1] Sick laser measurement systems technical description. Product Data Sheet, 2002.

[2] H. Andreasson, R. Triebel, and W. Burgard. Improving plane extraction from 3d data by fusing laser data and vision. In *IEEE IROS*, pages 2656–2661, August 2005.

[3] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[4] O. Jokinen. Self-calibration of a light striping system by matching multiple3-d profile maps. In *Second International Conference on 3-D Digital Imaging and Modeling*, pages 180–190, 1999.

[5] A. M. McIvor. Calibration of a laser stripe profiler. In *2nd International Conference on 3D Digital Imaging and Modeling (3DIM)*, pages 92–98, 1999.

[6] C. Mei and P. Rives. Calibration between a central catadioptric camera and a laser range finder for robotic applications. In *IEEE ICRA*, pages 532–537, May 2006.

[7] B. Tiddeman, M. Duffy, G. Rabey, and J. Lokier. Laser-video scanner calibration without the use of a frame store. *Vision, Image and Signal Processing*, 145(4):244–248, Aug 1998.

[8] G. Q. Wei and G. Hirzinger. Active self calibration of hand-mounted laser range finders. *IEEE Trans. Robotics and Automation*, 14(3):493–497, 1998.

[9] Q. Zhang and R. Pless. Constraints for heterogeneous sensor auto-calibration. In *IEEE Workshop on Realtime 3D Sensors and Their Use*, pages 38–43, 2004.

[10] Q. Zhang and R. Pless. Extrinsic calibration of a camera and laser range finder. In *IEEE IROS*, pages 2301–2306, 2004.

[11] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.

A Pose from Conic Section

We briefly describe a technique for estimating pose from a conic cross section. Let $A = RK$ and $B = A^T SA$; then (6) becomes

$$q^T A^T SA q = q^T B q = 0. \tag{11}$$

Note that the constraints are linear in the entries of matrix B. Since B is symmetric, and since it is defined only up to scale due to the homogeneity of (11), it has 5 degrees of freedom. A set of 5 linear constraints (i.e. 5 distinct points q) is therefore required for a unique solution, with each constraint of the form $c^T b = 0$, where

$$c^T = \begin{pmatrix} u^2 & 2uv & 2u & v^2 & 2v & 1 \end{pmatrix} \tag{12}$$

and b is a vector defining the relevant entries of the matrix B as

$$b^T = \begin{pmatrix} B_{11} & B_{12} & B_{13} & B_{22} & B_{23} & B_{33} \end{pmatrix}. \tag{13}$$

We form a constraint matrix C whose rows encode measured points (u_k, v_k) and take the form $c^T b = 0$, and solve the homogeneous system $Cb = 0$ by computing the eigenvector corresponding to the smallest eigenvalue of $C^T C$. We then factor the symmetric matrix B as $B = V\Lambda V^T$, where the columns of V are the eigenvectors and where Λ is a diagonal matrix of the eigenvalues. Using (11), we have

$$B = V\Lambda V^T = A^T SA, \tag{14}$$

so it follows that $A = \sqrt{S\Lambda}V^T$. The resulting matrix A may now be factored to obtain R and T as in Section 4.2. To refine these initial closed-form parameters, we again optimize using an iterative nonlinear algorithm; in particular, we estimate the pose parameters such that the sum of squared distances between RKq_k and the cone surface is minimized.