A Minimal Solution for the Extrinsic Calibration of a Camera and a Laser-Rangefinder

Francisco Vasconcelos, Joao P. Barreto, *Member*, and Urbano Nunes, *Senior Member*
E-mail: {fpv, jpbar, urbano}@isr.uc.pt

Abstract—This article presents a new algorithm for the extrinsic calibration of a perspective camera and an invisible 2D laser-rangefinder (LRF). The calibration is achieved by freely moving a checkerboard pattern in order to obtain plane poses in camera coordinates and depth readings in the LRF reference frame. The problem of estimating the rigid displacement between the two sensors is formulated as the one of registering a set of planes and lines in the 3D space. It is proved for the first time that the alignment of 3 plane-line correspondences has at most 8 solutions, that can be determined by solving a standard *p3p* problem and a linear system of equations. This leads to a minimal closed-form solution for the extrinsic calibration that can be used as hypothesis generator in a RANSAC paradigm. Our calibration approach is validated through simulation and real experiments, that show the superiority with respect to the current state-of-the-art method requiring a minimum of 5 input planes.

Index Terms—Extrinsic Calibration, Laser-Rangefinder, Euclidean Registration, Minimal Problems, Sensor Fusion.

---

1 INTRODUCTION

THERE are many systems and applications that combine perspective cameras and invisible 2D Laser-Rangefinder (LRF). A non-exhaustive list of examples includes the acquisition of ground-based city models by using an LRF for obtaining structure, and a camera for rendering texture [1]; the fusion of laser shape features with visual appearance for object classification [2] and pedestrian detection [3]; or the joint use of camera and laser for recognition and modeling of landmarks in outdoor self-localization and mapping [4]. In all these cases the fusion of the two sensor modalities requires knowing in advance the relative pose between camera and laser for projecting the depth readings into the images. Our article addresses this extrinsic calibration problem.

The number of published works in the extrinsic calibration of a camera and a LRF is relatively small. The most broadly used method was proposed by Zhang and Pless in [5], and describes a practical procedure where a checkerboard pattern is freely moved in front of the two sensors as shown in Fig. 1. The poses of the checkerboard are computed from plane-to-image homographies [6], and the camera coordinates of the planes are related with laser depth readings for establishing a set of linear constraints in the extrinsic calibration parameters. The solution of the system of equations provides an initial estimate for the relative rotation and translation, that is subsequently refined by iterative minimization of the re-projection error (bundle adjustment [7]). Zhang's algorithm suffers from two major drawbacks: (i) the system of linear equations does not directly enforce the
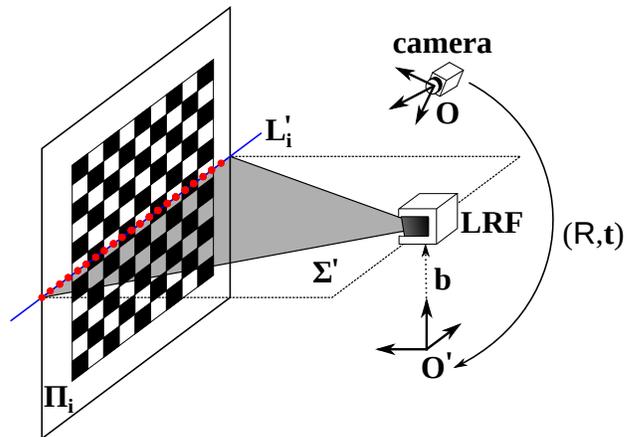

Fig. 1. Input data for estimating the relative pose between camera and LRF reference frames with origins in $\mathbf{O}$ and $\mathbf{O}'$. The checkerboard plane is represented in camera coordinates $\mathbf{\Pi}_i$ that are determined from the plane-image homography. The LRF reads the depth of the points lying on the line where the pattern intersects the scan plane $\mathbf{\Sigma}'$. The calibration problem is formulated as the registration of planes $\mathbf{\Pi}_i$ and co-planar lines $\mathbf{L}'_i$ that are fitted to the depth readings.

rotation matrix to be in *SO(3)*, which often leads to poor initialization that cause the iterative estimation to run into local minima; and (ii) the closed-form algorithm requires at least 5 input planes being clearly a non-minimal solution for the calibration problem.

This paper proposes a minimal solution for the described extrinsic calibration that estimates the rigid displacement between camera and LRF from 3 input planes. We fit lines to the laser depth readings and carry the

• *The authors are with the Institute for Systems and Robotics, Faculty of Science and Technology, University of Coimbra, Coimbra, Portugal.*

euclidean registration of 3 planes with 3 co-planar lines in an optimal and closed-form manner. It is shown that, by formulating the problem in the dual 3D space, the rotation and translation can be estimated separately. Finding the rotation is equivalent to determining the relative orientation between two views with known baseline. As proved in [8], this last problem can be cast as a standard *p3p* problem [9] admitting at most 8 distinct solutions. For each rotation solution there is a corresponding translation that can be found by solving an additional system of linear equations. Our main contribution is this new registration algorithm that is used as an efficient hypothesis generator in a RANSAC paradigm [10] for robust camera-LRF calibration. The minimal solution is tested in simulation and its singularities are discussed. Experiments using both synthetic and real data show that the proposed calibration method outperforms the state-of-the-art [5] in terms of robustness, accuracy, and required number of input planes. As a side result, we provide a more intuitive proof that the problem of estimating a relative rotation between two cameras with known baseline can be formulated as a *p3p* problem [8].

1.1 Related Work

This article is closely related with Zhang's work [5], where the extrinsic camera-LRF calibration is achieved by freely moving a checkerboard pattern. The procedure is simple to execute, and the checkerboard images can be used in parallel for calibrating the camera intrinsics [11], [12]. Like in [5], it is possible to use our method to jointly refine the intrinsic and extrinsic calibration during the final global optimization step. However, we do not discuss this feature and assume that the camera intrinsics are accurately known at all time. The calibration method presented in [13] is conceptually equivalent to [5], as is acknowledged by its authors.

Alternatives to Zhang's method can only be applied to a limited set of situations. Some contributions assume specific setups, like the LRF mounted on a calibrated rotating platform [14], or prior information, like an initial pose obtained through physical measurements [15], [16]. In other cases additional inertial data is used [17], [18]. A minimal solution for the extrinsic calibration of a camera and a LIght Detection And Ranging sensor (LIDAR) has been recently proposed [19]. The method uses a planar pattern for establishing correspondences between points in the LIDAR and lines in the image. The calibration problem is formulated as the 3D registration of co-planar points with planes intersecting into a single point. This leads to a system of polynomial equations that is solved using Macaulay resultants, obtained from 6 input images. Another minimal solution was proposed to calibrate both intrinsic and extrinsic parameters of a camera and a visible range finder [20]. In this case it is easy to make data associations between laser depth readings and visible laser dots projected on the camera without using a calibration target. This procedure requires the LRF-camera system to acquire 3 dot associations in 5 different positions. These last two approaches cannot be directly extended to the invisible LRF because they also use additional sensory information for the data association.

Since we formulate the camera-LRF calibration as the problem of aligning planes with co-planar lines, the article also relates with the literature in 3D euclidean registration and related topics. In particular we use previous results in registering two clouds of 3 or more 3D points [21]; in estimating the camera pose from the images of 3 or more 3D points (the so called Perspective-n-Pose (*PnP*) problem) [9], [22]; and in determining from 3 correspondences the relative rotation between two views with known baseline [8]. Olsson et al. have recently proposed in [23] a Branch-and-Bound framework to solve different euclidean registration problems: point-to-point, point-to-line, and point-to-plane. Within this topic, the recent work of Ramalingam et al. [24] in minimal solutions for the registration of points and planes is specially relevant. It is possible to adapt their algorithm for aligning 3 planes with 3 generic lines where each line is parametrized as a pair of points. Although such approach can eventually lead to a minimal solution for camera-LRF calibration, we propose an alternative registration algorithm that simplifies the problem by conveniently exploring the fact that the lines are co-planar.

1.2 Notation

Scalars are represented by plain letters, e.g. λ, vectors are indicated by bold symbols, e.g. $\mathbf{n}$, $\mathbf{O}$, and matrices are denoted by letters in sans serif font, e.g. T. We do not distinguish between a linear transformation and a matrix representing it. Planes are represented by an homogeneous vector with dimension 4 that is indicated by an uppercase greek letter, e.g. $\mathbf{\Pi}$, and 3D lines are expressed in homogeneous Plucker coordinates, e.g. the 6×1 vector $\mathbf{L}$. The equality up to scale is denoted by $\sim$ in order to be distinguished from the strict equality $=$. Vector cross product and L2 norm are indicated by $\times$ and $\|\ \|$ respectively. We use a prime symbol to indicate geometric entities represented in the laser coordinate system, e.g. $\mathbf{\Pi}'$.

2 THE CALIBRATION PROBLEM

Consider a camera and a LRF for which the local coordinate systems have origin in $\mathbf{O}$ and $\mathbf{O}'$ as shown in Fig. 1. The extrinsic calibration aims to determine the rigid transformation T such that:

$$\begin{pmatrix} \mathbf{Q}' \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} \mathsf{R} & \mathbf{t} \\ \mathbf{0}_3^\mathsf{T} & 1 \end{pmatrix}}_{\mathsf{T}} \begin{pmatrix} \mathbf{Q} \\ 1 \end{pmatrix}, \qquad (1)$$

where $\mathbf{Q}$ and $\mathbf{Q}'$ are respectively non-homogeneous point coordinates in camera and LRF reference frames, R denotes a rotation matrix, and $\mathbf{t}$ is the translation vector.

In [5] the calibration is carried from N images of a checkerboard pattern that is freely moved in front of the two sensors. Let $\mathbf{\Pi}_i$ be the homogeneous representation of the calibration plane in camera coordinates, that is estimated from plane-to-image point correspondences [6]:

$$\mathbf{\Pi}_i \sim \begin{pmatrix} \mathbf{n}_i \\ 1 \end{pmatrix}, \qquad i = 1, 2 \dots N \qquad (2)$$

For each plane $\mathbf{\Pi}_i$, the LRF provides depth readings of a set of 3D points $\mathbf{Q}'_{ij}$ that lie on the line where the checkerboard intersects the scan plane $\mathbf{\Sigma}'$. Let the non-homogeneous coordinates in the laser reference frame be

$$\mathbf{Q}'_{ik} = \begin{pmatrix} x_{ik} \\ y_{ik} \\ z_{ik} \end{pmatrix}, \qquad k = 1, 2 \dots K_i$$

Zhang and Pless assume, without loss of generality, that $\mathbf{\Sigma}'$ is coincident with the Y plane. By inverting equation 1 and taking into account that y_{ik} is always zero, it follows that:

$$\mathbf{Q}_{ik} = \mathsf{R}^\mathsf{T} \underbrace{\begin{pmatrix} 1 & 0 & \\ 0 & 0 & -\mathbf{t} \\ 0 & 1 & \end{pmatrix}}_{\mathsf{H}} \mathring{\mathbf{Q}}_{ik}, \qquad (3)$$

with $\mathbf{Q}_{ik}$ being the point representation in camera coordinates and

$$\mathring{\mathbf{Q}}_{ik} = \begin{pmatrix} x_{ik} \\ z_{ik} \\ 1 \end{pmatrix}.$$

Since the points detected by the laser are in the checkerboard pattern, then the following must hold

$$\mathbf{\Pi}_i^\mathsf{T} \mathbf{Q}_{ik} = 0.$$

Replacing by the results of equations 2 and 3, it yields that

$$\mathbf{n}_i^\mathsf{T} \mathsf{H} \mathring{\mathbf{Q}}_{ik} = -1, \quad \forall_{i,k} \qquad (4)$$

In summary, the checkerboard planes $\mathbf{\Pi}_i$, expressed in camera coordinates, and the points $\mathbf{Q}'_{ik}$, represented in LRF coordinates, define a set of linear constraints in the entries of matrix H that encodes the rigid displacement between the two sensors. In [5], Zhang and Pless propose to compute H in a DLT-like manner, and factorize the result into the rotation R and translation $\mathbf{t}$. Unfortunately the linear estimation of matrix H is carried without enforcing the structure of equation 3. This means that in general the direct factorization does not provide a valid rotation matrix R, and a non-optimal projection into *SO(3)* is required. Moreover, for each calibration plane $\mathbf{\Pi}_i$ there are only two constraints of the form of equation 4 that are linearly independent. Since matrix H has 9 entries, then the estimation requires $N \geq 5$ calibration planes. The fact that the calibration algorithm is non-minimal and sub-optimal often leads to erroneous results as shown later in the paper.

3 REGISTRATION IN THE DUAL SPACE

While in [5] the extrinsic calibration is achieved by finding the transformation T that aligns points $\mathbf{Q}'_{ik}$ with planes $\mathbf{\Pi}_i$, in here we propose to fit lines to the laser points and formulate the problem as the 3D registration of a set of co-planar lines $\mathbf{L}'_i$ with a set of planes $\mathbf{\Pi}_i$. In other words, we aim to find the rotation R and the translation $\mathbf{t}$ such that the planes $\mathbf{\Pi}'_i$, given by

$$\mathbf{\Pi}'_i = \underbrace{\begin{pmatrix} \mathsf{R} & \mathbf{0} \\ -\mathbf{t}^\mathsf{T}\mathsf{R} & 1 \end{pmatrix}}_{\mathsf{T}^{-\mathsf{T}}} \mathbf{\Pi}_i, \qquad i = 1, 2 \dots N \qquad (5)$$

go through the lines $\mathbf{L}'_i$.

It is well known that points and planes are dual entities in 3D, with a plane in the projective space $\mathcal{P}^3$ being represented as a point in the dual space $\mathcal{P}^{3^*}$ and vice-versa. Thus, and as shown in Fig. 2(a), equation 5 can be understood as a projective transformation in $\mathcal{P}^{3^*}$ that maps points $\mathbf{\Pi}_i$ into points $\mathbf{\Pi}'_i$. Henceforth, and if nothing is said, we will reason in terms of the dual space for deriving the desired registration algorithm.

The dual of a line $\mathbf{L}'_i$ in $\mathcal{P}^3$ is a line $\mathbf{L}^*_i$ in $\mathcal{P}^{3^*}$ whose representation in Plücker coordinates is

$$\mathbf{L}^*_i \sim \begin{pmatrix} \mathbf{v}'_i \\ \mathbf{u}'_i \end{pmatrix}, \qquad i = 1, 2 \dots N \qquad (6)$$

where the 3-dimension vectors $\mathbf{u}'_i$ and $\mathbf{v}'_i$ denote the direction and momentum of the original line $\mathbf{L}'_i$ [25]. A generic point $\mathbf{\Pi}'$ on $\mathbf{L}^*_i$ is the dual representation of a plane that contains the line $\mathbf{L}'_i$. Moreover, and since the lines $\mathbf{L}'_i$ that are detected by the LRF are necessarily coplanar, it comes that the dual lines $\mathbf{L}^*_i$ must intersect into a point $\mathbf{\Sigma}'$ that represents the laser scan plane.

The extrinsic calibration consists in finding the transformation $\mathsf{T}^{-\mathsf{T}}$ in the dual space that maps points $\mathbf{\Pi}_i$, expressed in camera coordinates, into points $\mathbf{\Pi}'_i$ lying in the lines $\mathbf{L}^*_i$ detected by the LRF (see Fig. 2(a)). The latter points are obviously the dual representation of the calibration planes expressed in laser coordinates. The transformation $\mathsf{T}^{-\mathsf{T}}$, provided in equation 5, can be factorized in two distinct transformations:

1) A rotation transformation M, that maps points $\mathbf{\Pi}_i$ into points $\widehat{\mathbf{\Pi}}_i$ as show in Fig. 2(a):

$$\widehat{\mathbf{\Pi}}_i \sim \begin{pmatrix} \mathsf{R} & \mathbf{0} \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix} \mathbf{\Pi}_i \qquad (7)$$

This rotation must be such that each line $\mathbf{W}_i$, defined by $\widehat{\mathbf{\Pi}}_i$ and the origin of the dual space $\mathbf{O}^*$, intersects the corresponding line $\mathbf{L}^*_i$.

2) A transformation S that moves the points $\widehat{\mathbf{\Pi}}_i$ along the lines $\mathbf{W}_i$ in order to map them into points $\mathbf{\Pi}'_i$:

$$\mathbf{\Pi}'_i \sim \begin{pmatrix} \mathsf{I}_{3\times3} & \mathbf{0}_3 \\ -\mathbf{t}^\mathsf{T} & 1 \end{pmatrix} \widehat{\mathbf{\Pi}}_i \qquad (8)$$

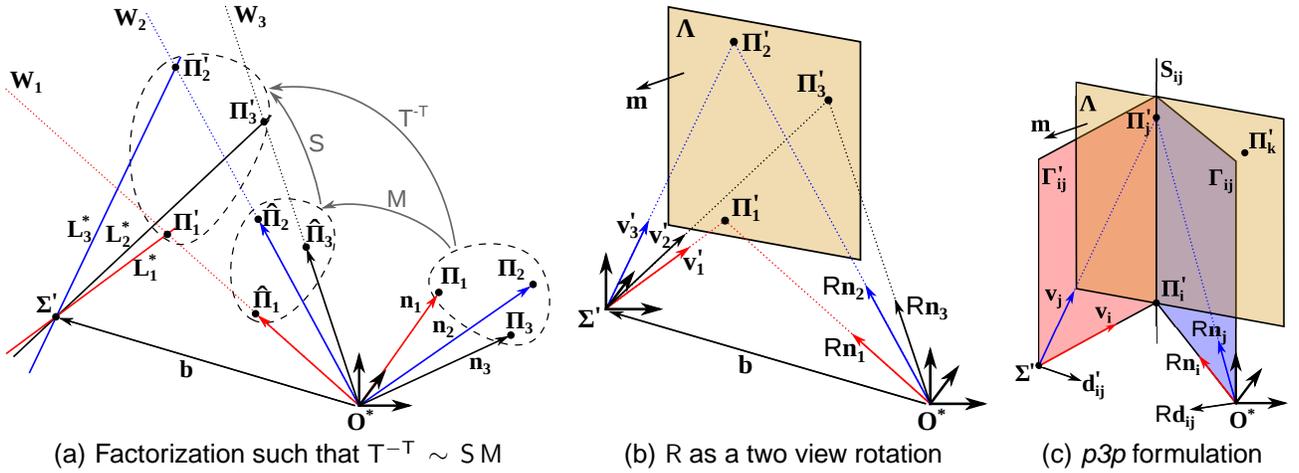with $\mathsf{I}_{3\times3}$ being the 3×3 identity matrix.

(a) Factorization such that $\mathsf{T}^{-\top} \sim \mathsf{S}\,\mathsf{M}$ (b) R as a two view rotation (c) *p3p* formulation

Fig. 2. Reasoning in the dual projective space $\mathcal{P}^{3^*}$ for estimating the transformation $\mathsf{T}^{-\top}$ that aligns $N = 3$ planes $\mathbf{\Pi}_i$ with $N = 3$ co-planar lines $\mathbf{L}'_i$: (a) $\mathsf{T}^{-\top}$ can be factorized into a rotation M, such that the rotated lines $\mathbf{W}_i$ intersect the lines $\mathbf{L}^*_i$, and a shift transformation S, that makes the rotated points $\widehat{\mathbf{\Pi}}_i$ to be coincident with the intersection points $\mathbf{\Pi}'_i$; (b) Computing the rotation is equivalent to estimating the relative orientation between two views with known baseline $\mathbf{b}$ (we think on the systems of coordinates with center in $\mathbf{O}^*$ and $\mathbf{\Sigma}'$ as being parallel to each other); (c) The estimation of R is formulated as a *p3p* problem by observing that planes $\mathbf{\Gamma}$, $\mathbf{\Gamma}'$ and $\mathbf{\Lambda}$ are in the same pencil.

The next two subsections show how M and S can be estimated from $N = 3$ correspondences between planes and co-planar lines.

3.1 The Rotation Transformation

Since the lines $\mathbf{L}^*_i$ are the dual representation of co-planar lines, then they define a pencil in $\mathcal{P}^{3^*}$ that goes through $\mathbf{\Sigma}'$. We assume, without loss of generality, that the origin of the LRF reference frame never lies in the scan plane, and that $\mathbf{\Sigma}'$ has the following projective representation:

$$\mathbf{\Sigma}' \sim \begin{pmatrix} \mathbf{b} \\ 1 \end{pmatrix}. \tag{9}$$

A second pencil of lines can be obtained by joining the dual points $\mathbf{\Pi}_i$ with the origin of the coordinate system $\mathbf{O}^*$ (see Fig. 2(a)). The objective is to find the rotation around an axis passing by $\mathbf{O}^*$ that makes the second pencil to intersect the first.

3.1.1 R *as the relative rotation between two virtual views*

We can think on the two line pencils of Fig. 2(a) as being two pin-hole cameras with projection centers in $\mathbf{\Sigma}'$ and $\mathbf{O}^*$. By doing so the problem of finding the rotation R such that lines $\mathbf{W}_i$ intersect lines $\mathbf{L}^*_i$ becomes geometrically equivalent to determining the relative orientation between two virtual views when the baseline direction is known (Fig. 2(b)). Moriya and Takeda solved this problem in [8] by showing how to compute R from $N = 3$ point correspondences between images.

In our case the projective coordinates of the image points in the virtual view $\mathbf{O}^*$ are given by vectors $\mathbf{n}_i$. These vectors are the normal directions of the calibration planes $\mathbf{\Pi}_i$ provided in equation 2. Likewise, the image

points in the virtual view $\mathbf{\Sigma}'$ are given by vectors $\mathbf{v}'_i$, that correspond to the momenta of the lines $\mathbf{L}'_i$ detected by the LRF (equation 6). The baseline direction is defined by the projective coordinates of the laser scan plane that are provided in equation 9. This geometric construction is illustrated in Fig. 2(b), where the local reference frame with center in $\mathbf{\Sigma}'$ is parallel to the world system of coordinates with origin in $\mathbf{O}^*$.

The next section provides an alternative and more intuitive derivation of the solution proposed in [8], and shows how the relative orientation R can be determined from $N = 3$ correspondences $\mathbf{n}_i$, $\mathbf{v}'_i$ that are interpreted as image points in two virtual views separated by a baseline with direction $\mathbf{b}$.

3.1.2 *Determining* R *by solving a P3P problem*

If the relative orientation R is known, then the back-projection lines with directions $\mathbf{v}'_i$ and $\mathsf{R}\,\mathbf{n}_i$ intersect into a 3D point $\mathbf{\Pi}'_i$ as shown in Fig. 2(b). Thus, if we consider the $N = 3$ correspondences, we obtain 3 points that uniquely define a plane. Let $\mathbf{\Lambda}$ be this plane that is expressed in world coordinates by:

$$\mathbf{\Lambda} \sim \begin{pmatrix} \mathbf{m} \\ 1 \end{pmatrix}. \tag{10}$$

Consider now two of these points, $\mathbf{\Pi}'_i$ and $\mathbf{\Pi}'_j$, that define a 3D line $\mathbf{S}_{ij}$ as shown in Fig. 2(c). Let $\mathbf{\Gamma}_{ij}$ be the plane containing both $\mathbf{S}_{ij}$ and the origin $\mathbf{O}^*$. It is easy to see that the normal direction to $\mathbf{\Gamma}_{ij}$ must be $\mathsf{R}\,\mathbf{d}_{ij}$ with

$$\mathbf{d}_{ij} \sim \mathbf{n}_i \times \mathbf{n}_j. \tag{11}$$

Taking into account that the plane contains the $\mathbf{O}^*$, it follows that its projective representation in world

coordinates is

$$\mathbf{\Gamma}_{ij} \sim \begin{pmatrix} \mathsf{R}\mathbf{d}_{ij} \\ 0 \end{pmatrix} \qquad (12)$$

In a similar manner, the line $\mathbf{S}_{ij}$ and point $\mathbf{\Sigma}'$ define a plane $\mathbf{\Gamma}'_{ij}$ with normal

$$\mathbf{d}'_{ij} \sim \mathbf{v}'_i \times \mathbf{v}'_j \qquad (13)$$

From equation 9, it comes that the projective representation of $\mathbf{\Gamma}'_{ij}$ in world coordinates is :

$$\mathbf{\Gamma}'_{ij} \sim \begin{pmatrix} \mathbf{d}'_{ij} \\ -\mathbf{b}^{\mathsf{T}}\mathbf{d}'_{ij} \end{pmatrix} \qquad (14)$$

The line $\mathbf{S}_{ij}$ defines a 1-D pencil of planes that contains, not only $\mathbf{\Gamma}_{ij}$ and $\mathbf{\Gamma}'_{ij}$, but also the plane $\mathbf{\Lambda}$ that passes by the three reconstructed points. Thus, and since $\mathbf{\Gamma}_{ij}$ and $\mathbf{\Gamma}'_{ij}$ are always distinct planes that can be used as a projective basis for the pencil, there is a pair of scalars α_{ij} and β_{ij} such that the following holds:

$$\mathbf{\Lambda} \sim \alpha_{ij}\mathbf{\Gamma}_{ij} + \beta_{ij}\mathbf{\Gamma}'_{ij}$$

Replacing $\mathbf{\Lambda}$, $\mathbf{\Gamma}_{ij}$, and $\mathbf{\Gamma}'_{ij}$ by the results of equations 10, 12, and 14, yields:

$$\begin{pmatrix} \mathbf{m} \\ 1 \end{pmatrix} \sim \begin{pmatrix} \alpha_{ij}\,\mathsf{R}\,\mathbf{d}_{ij} + \beta_{ij}\,\mathbf{d}'_{ij} \\ -\beta_{ij}\mathbf{b}^{\mathsf{T}}\,\mathbf{d}'_{ij} \end{pmatrix} \qquad (15)$$

We can fixate the scale factor by making

$$\beta_{ij} = -\frac{1}{\mathbf{b}^{\mathsf{T}}\mathbf{d}'_{ij}}.$$

Replacing β_{ij} in equation 15 leads to the strict equality

$$\mathbf{m} = -\frac{1}{\mathbf{b}^{\mathsf{T}}\mathbf{d}'_{ij}}\,\mathbf{d}'_{ij} + \alpha_{ij}\,\mathsf{R}\,\mathbf{d}_{ij}\,,$$

that can be re-written as

$$\alpha_{ij}\,\mathbf{d}_{ij} = \mathsf{R}^{\mathsf{T}}(\frac{1}{\mathbf{b}^{\mathsf{T}}\mathbf{d}'_{ij}}\mathbf{d}'_{ij} + \mathbf{m})\,.$$

Since the $N = 3$ correspondences $\mathbf{n}_i$, $\mathbf{v}'_i$ give rise to three distinct equations, we can obtain the following system:

$$\begin{cases} \alpha_{12}\,\mathbf{d}_{12} = \mathsf{R}^{\mathsf{T}}(\mathbf{P}'_{12} + \mathbf{m}) \\ \alpha_{13}\,\mathbf{d}_{13} = \mathsf{R}^{\mathsf{T}}(\mathbf{P}'_{13} + \mathbf{m}) \\ \alpha_{23}\,\mathbf{d}_{23} = \mathsf{R}^{\mathsf{T}}(\mathbf{P}'_{23} + \mathbf{m}) \end{cases} \qquad (16)$$

with

$$\mathbf{P}'_{ij} = \frac{1}{\mathbf{b}^{\mathsf{T}}\mathbf{d}'_{ij}}\mathbf{d}'_{ij}. \qquad (17)$$

A careful analysis shows that the rotation R can be found by considering two well known geometric problems: the unknowns α_{ij} can be treated as scalar depths to be determined as solutions of a perspective-3-pose (*p3p*) problem [9]; and, given the non-homogeneous representation of point correspondences $\alpha_{ij}\mathbf{d}_{ij}$, $\mathbf{P}'_{ij}$, the rotation R and the shift $\mathbf{m}$ can be computed by solving for the relative orientation between reference frames
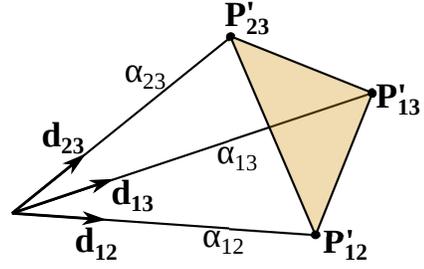

Fig. 3. The classical *p3p* problem [9]: given the distances between points $\mathbf{P}'_{12}$, $\mathbf{P}'_{13}$, and $\mathbf{P}'_{23}$, and the angles between vectors $\mathbf{d}_{12}$, $\mathbf{d}_{13}$, and $\mathbf{d}_{23}$, determine the unknown depths α_{12}, α_{13}, and α_{23}.

[21]. Re-arranging the system of equation 16 in order to eliminate R and $\mathbf{m}$, it yields

$$\begin{cases} ||\alpha_{12}\mathbf{d}_{12} - \alpha_{13}\mathbf{d}_{13}|| = ||\mathbf{P}'_{12} - \mathbf{P}'_{13}|| \\ ||\alpha_{12}\mathbf{d}_{12} - \alpha_{23}\mathbf{d}_{23}|| = ||\mathbf{P}'_{12} - \mathbf{P}'_{23}|| \\ ||\alpha_{13}\mathbf{d}_{13} - \alpha_{23}\mathbf{d}_{23}|| = ||\mathbf{P}'_{13} - \mathbf{P}'_{23}|| \end{cases} . \qquad (18)$$

Consider, without loss of generality. that the vectors $\mathbf{d}_{ij}$ have unitary norm (equation 11). By applying the law of cosines, we arrive at the classic equations of the *p3p* problem:

$$\begin{cases} \alpha_{12}^2 + \alpha_{13}^2 - \alpha_{12}\alpha_{12}\mathbf{d}_{12}{}^{\mathsf{T}}\mathbf{d}_{13} = ||\mathbf{P}'_{12} - \mathbf{P}'_{13}||^2 \\ \alpha_{12}^2 + \alpha_{23}^2 - \alpha_{12}\alpha_{12}\mathbf{d}_{12}{}^{\mathsf{T}}\mathbf{d}_{23} = ||\mathbf{P}'_{12} - \mathbf{P}'_{23}||^2 \\ \alpha_{13}^2 + \alpha_{23}^2 - \alpha_{12}\alpha_{12}\mathbf{d}_{13}{}^{\mathsf{T}}\mathbf{d}_{23} = ||\mathbf{P}'_{13} - \mathbf{P}'_{23}||^2 \end{cases} . \qquad (19)$$

As shown in Fig. 3, the scalars α_{12}, α_{13}, and α_{23} are understood as unknown depth values that can be determined by applying any of the algorithms described in [9] (in our experiments we will consider Grunert's method [26]). By substituting the depth values back in equation 16, it yields a system of equations relating the non-homogeneous coordinates of 3 points expressed in different reference frames. The euclidean transformation R and $\mathbf{m}$ between these two frames can be computed in a straightforward manner by applying the absolute orientation algorithm proposed in [21].

3.1.3 Discussion

Let's better understand how the system of equations 16 relates with the original problem of registering planes with co-planar lines. From equation 11 it comes that $\mathbf{d}_{ij}$ is the cross-product of the normals of $\mathbf{\Pi}_i$ and $\mathbf{\Pi}_j$, meaning that $\mathbf{d}_{ij}$ is the direction in camera coordinates of the line where the two planes meet. On the other hand $\mathbf{d}'_{ij}$ is the cross-product of the momenta of lines $\mathbf{L}'_i$ and $\mathbf{L}'_j$, meaning that it corresponds to the direction of the line that is defined by the origin $\mathbf{O}'$ and the point where $\mathbf{L}'_i$ and $\mathbf{L}'_j$ intersect. From equation 9 it comes in a straightforward manner that the 3-vector $\mathbf{P}'_{ij}$ provided in equation 17 is the non-homogeneous representation in laser coordinates of the point intersection of lines $\mathbf{L}'_i$
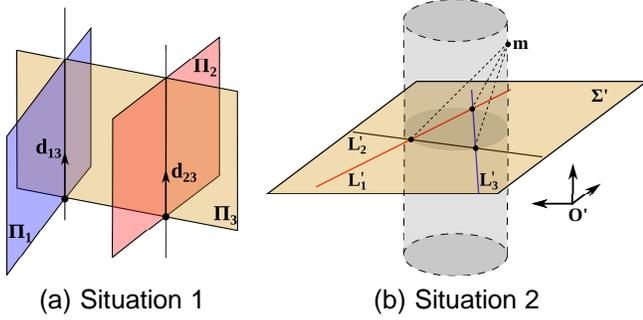
(a) Situation 1   (b) Situation 2

Fig. 4. Singular configurations while determining the relative rotation R. Situation 1: The lines where the checkerboard planes intersect are parallel. Situation 2: The 3 checkerboard planes intersect into a point $\mathbf{m}$ that lies in the *danger cylinder* [9] defined by the intersection points of the lines detected by the LRF.

and $\mathbf{L}'_j$. Finally, and as shown by Fig. 2(b), the vector $\mathbf{m}$ denotes the normal to the plane that is defined by $\mathbf{\Pi}'_1$, $\mathbf{\Pi}'_2$, and $\mathbf{\Pi}'_3$ in the dual space. Translating to $\mathcal{P}^3$ it comes that $\mathbf{m}$ is the non-homogeneous representation in the LRF reference frame of the point where the 3 calibration planes intersect.

From the above it follows that the *p3p* formulation of equation 16 can be understood as the estimation of the pose of a virtual pine-hole camera with projection center in $\mathbf{m}$, that observes the vertices of the triangle defined by the lines $\mathbf{L}'_1$, $\mathbf{L}'_2$, and $\mathbf{L}'_3$ (see Fig. 4(b)). Remark that determining $\mathbf{m}$ is not relevant for our plane-line registration problem. Thus, and after finding the scalar depths using Grunert's algorithm [26], the absolute orientation is only solved for the rotation unknown, meaning that the translation component $\mathbf{m}$ in equation 16 is not explicitly computed. It is well known that the standard *p3p* problem has up to 4 distinct real solutions [9]. Unfortunately the scalar unknowns α_{ij} in equation 16 do not have the physical meaning of depth and are allowed to take negative values. Hence there can be up to 8 distinct solutions for the *p3p* problem of equation 19 leading to the same number of relative rotations R.

It is well known that the *p3p* problem degenerates whenever one of the following situations arise: (i) the camera center is at the infinity or is in the same plane as the 3 control points; and (ii) the camera center is in the cylinder that contains the 3 control points and is orthogonal to the plane defined by them (the *danger cylinder*) [9]. In the context of our problem the first situation occurs whenever one of the checkerboard planes is parallel to (or contains) the line where the two other planes intersect (Fig. 4(a)). The second case arises whenever the point where the 3 checkerboard planes meet lies in the danger cylinder defined by the intersections of the lines reconstructed by the LRF (Fig. 4(b)).

3.2 The Projective Scaling Transformation

From the previous section we obtain up to 8 solutions for the rotation transformation M. For each possible M, the dual points $\mathbf{\Pi}_i$ are mapped into $\widehat{\mathbf{\Pi}}_i$ according to equation 7, and the lines $\mathbf{W}_i$ are determined by joining $\widehat{\mathbf{\Pi}}_i$ with the origin $\mathbf{O}^*$ as shown in Fig. 2(a). The representation $\mathbf{\Pi}'_i$ of the calibration planes in the LRF reference frame are determined by intersecting corresponding lines $\mathbf{L}^*_i$ and $\mathbf{W}_i$ [27]. The next step is to compute the projective scaling S that maps $\widehat{\mathbf{\Pi}}_i$ into $\mathbf{\Pi}'_i$, and defines in a unique manner the relative translation $\mathbf{t}$ between the two sensors. From the stated, we start by replacing $\mathbf{\Pi}'_i$ and $\widehat{\mathbf{\Pi}}_i$ in equation 8 yielding

$$\mu_i \begin{pmatrix} \mathbf{n}'_i \\ 1 \end{pmatrix} = \begin{pmatrix} \mathsf{I}_{3\times3} & \mathbf{0}_3 \\ -\mathbf{t}^\mathsf{T} & 1 \end{pmatrix} \begin{pmatrix} \mathsf{R}\mathbf{n}_i \\ 1 \end{pmatrix}$$

where $\mathbf{n}'_i$ is the normal to the plane $\mathbf{\Pi}'_i$ and μ_i denotes an unknown scale factor. Considering the top three equations, it comes that

$$\mu_i = \frac{\mathbf{n}'^\mathsf{T}_i \, \mathsf{R} \, \mathbf{n}_i}{\mathbf{n}'^\mathsf{T}_i \, \mathbf{n}'_i}$$

and replacing μ_i in the bottom equation leads to

$$\mathbf{n}'^\mathsf{T}_i \, \mathbf{n}'_i \, \mathbf{n}_i^\mathsf{T} \, \mathsf{R}^\mathsf{T} \, \mathbf{t} - \mathbf{n}'^\mathsf{T}_i \, \mathbf{n}'_i + \mathbf{n}'^\mathsf{T}_i \, \mathsf{R} \, \mathbf{n}_i = 0 \,.$$

Thus, each calibration plane gives rise to linear constraint in the entries of the vector $\mathbf{t}$, which means that the translation can be computed in a straightforward manner by making

$$\mathbf{t} = \mathsf{A}^{-1}\mathbf{b} \tag{20}$$

with

$$\mathsf{A} = \begin{pmatrix} \mathbf{n}'^\mathsf{T}_1 \, \mathbf{n}'_1 \, \mathbf{n}_1^\mathsf{T} \\ \mathbf{n}'^\mathsf{T}_2 \, \mathbf{n}'_2 \, \mathbf{n}_2^\mathsf{T} \\ \mathbf{n}'^\mathsf{T}_3 \, \mathbf{n}'_3 \, \mathbf{n}_3^\mathsf{T} \end{pmatrix} \mathsf{R}^\mathsf{T},$$

and

$$\mathbf{b} = \begin{pmatrix} \mathbf{n}'^\mathsf{T}_1 \, \mathbf{n}'_1 - \mathbf{n}_1^\mathsf{T}\mathsf{R}\mathbf{n}_1 \\ \mathbf{n}'^\mathsf{T}_2 \, \mathbf{n}'_2 - \mathbf{n}_2^\mathsf{T}\mathsf{R}\mathbf{n}_2 \\ \mathbf{n}'^\mathsf{T}_3 \, \mathbf{n}'_3 - \mathbf{n}_3^\mathsf{T}\mathsf{R}\mathbf{n}_3 \end{pmatrix}.$$

In summary, for each possible rotation R satisfying the system of equations 16, there is a corresponding translation $\mathbf{t}$ that can be determined by equation 20. Remark that the 3×3 matrix A is singular *iff* two calibration planes are parallel to each other. This is the only case for which it is not possible to find a solution for the translation. However, and since this situation is a particular instance of the degenerate configuration of Fig. 4(a), the determination of $\mathbf{t}$ does not introduce new singularities in the proposed registration algorithm.

3.3 Outline of the registration algorithm

The algorithm for aligning $N = 3$ planes $\mathbf{\Pi}_i$ with the same number of co-planar lines $\mathbf{L}'_i$ can be summarized as follows:

1) For each two planes $\mathbf{\Pi}_i$, $\mathbf{\Pi}_j$ determine $\mathbf{d}_{ij}$ by cross-multiplying their normals (equation 11).
2) For each two lines $\mathbf{L}'_i$, $\mathbf{L}'_j$ determine the vector $\mathbf{d}'_{ij}$ by cross-multiplying their momenta (equation 13).
3) Determine the orthogonal direction $\mathbf{b}$ to the plane $\mathbf{\Sigma}'$ that contains the lines $\mathbf{L}'_i$ (equation 9).
4) Given $\mathbf{b}$, $\mathbf{d}_{ij}$, and $\mathbf{d}'_{ij}$, with $ij = 12, 13, 23$, formulate the *p3p* problem of equation 16.
5) Solve the *p3p* problem with respect to the rotation R using any standard approach [9]. There are $M \leq 8$ distinct solutions.
6) For each possible R, compute $\mathbf{\Pi}'_i$ by intersecting the the dual of the input line $\mathbf{L}'_i$ with the rotated line $\mathbf{W}_i$ as is shown in Fig. 2(a) (for the intersection of lines in 3D see section 3.3. in [28]).
7) Given R, $\mathbf{\Pi}_i$, and $\mathbf{\Pi}'_i$, with $i = 1, 2, 3$, apply equation equation 20 for determining the translation $\mathbf{t}$.

4 EXTRINSIC CALIBRATION ALGORITHM

Section 3 derives a closed-form algorithm for computing the $M \leq 8$ rigid transformations that align 3 planes with 3 co-planar lines. We now show how this new registration method can be used for obtaining the extrinsic calibration between the camera and the laser.

Let's recall that the inputs for calibration are the planes $\mathbf{\Pi}_i$ with $i = 1, 2 \ldots N$, expressed in camera coordinates, and the points $\mathbf{Q}'_{ik}$ with $k = 1, 2 \ldots K_i$, represented in non-homogeneous LRF coordinates. The application of plane-line registration requires fitting lines $\mathbf{L}'_i$ to the points $\mathbf{Q}'_{ik}$ using a standard regression method. If the number of input planes is $N = 3$, then the registration algorithm provides $M \leq 8$ solutions $\mathsf{T}^{(m)}$ with $m = 1, 2 \ldots M$, but we cannot decide about the rigid displacement between the two sensors without further information. For the case of $N > 3$, each triplet of plane-line correspondences gives rise to a set of solutions, and the correct relative pose T can be found using an hypothesize-and-test framework as detailed in section 4.1. In both situations the final calibration estimates can be further refined by minimizing the re-projection errors in the camera and LRF using iterative non-linear optimization. This bundle adjustment step is discussed in section 4.2.

4.1 Initial Estimation

Consider $N > 3$ correspondences between planes $\mathbf{\Pi}_i$ and lines $\mathbf{L}'_i$. The initial estimate T for the extrinsic calibration is obtained as follows:

1) Select 3 correspondences and apply the algorithm of section 3.3 for finding the transformations $\mathsf{T}^{(m)}$ that align lines and planes ($m = 1, 2 \ldots M$).
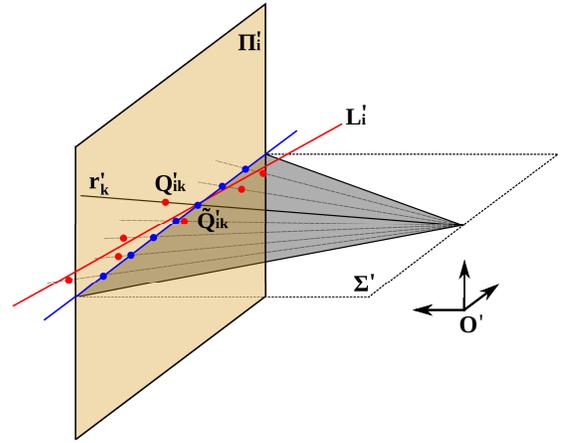

Fig. 5. The LRF reconstructs points $\mathbf{Q}'_{ik}$ (red) by measuring depth along the radial directions $\mathbf{r}'_k$. The minimization is carried over the distance between points $\mathbf{Q}'_{ik}$ and points $\widetilde{\mathbf{Q}}'_{ik}$ (blue). The latter are obtained by mapping the inlier planes $\mathbf{\Pi}_i$ into the LRF reference frame and intersecting the result with the radial directions $\mathbf{r}'_k$.

2) For each solution $\mathsf{T}^{(m)}$, compute the LRF coordinates $\mathbf{\Pi}_j'^{(m)}$ of the remaining $N - 3$ planes, and determine the euclidean distance $d_j^{(m)}$ in the dual space between $\mathbf{\Pi}_j'^{(m)}$ and the corresponding line $\mathbf{L}'_j$.
3) Rank each solution $\mathsf{T}^{(m)}$ by assigning the score

$$rank(\mathsf{T}^{(m)}) = \sum_j \max(t, d_j^{(m)}),$$

where t is a pre-defined threshold. This operator is similar to the one used in the MSAC robust estimator proposed in [29].
4) If $rank(\mathsf{T}) > rank(\mathsf{T}^{(m)})$, then make $\mathsf{T} = \mathsf{T}^{(m)}$ and consider as inliers the correspondences for which $d_j^{(m)} < t$ (the 3 correspondences that generated the solution have $d_j^{(m)} = 0$).
5) Return to step 1 for a new iteration.

Since the number of input correspondences is usually small ($N < 20$), we run an exhaustive search where all possible plane-line triplets are considered as solution generators. For a large N the hypothesize-and-test can be performed in a Random Sample Consensus manner in order to keep the computation tractable [10].

4.2 Bundle Adjustment

The initialization procedure provides an extrinsic calibration T and a set of plane-line correspondences that are classified as inliers. The calibration accuracy can be further improved by minimizing the re-projection errors in the camera and/or LRF using bundle adjustment [7].

The LRF measures depth along a set of radial directions $\mathbf{r}'_k$ that are uniformly distributed in the scan plane $\mathbf{\Sigma}'$ around the projection center. Please note that we assume that this projection center is not coincident

with the origin $\mathbf{O}'$ of the laser reference frame. The depth readings enable to reconstruct the points $\mathbf{Q}'_{ik}$ that give rise to the lines $\mathbf{L}'_i$ that are considered in the 3D registration. After obtaining an initial calibration estimate T, each inlier plane $\mathbf{\Pi}_i$ is mapped into the LRF reference frame using equation 5, and the resulting plane $\mathbf{\Pi}'_i$, expressed in laser coordinates, is intersected with the radial lines $\mathbf{r}'_k$ yielding a set of points $\widetilde{\mathbf{Q}}'_{ik}$ (see Fig. 5). The LRF residue to be minimized is the sum of the square distances between the points $\widetilde{\mathbf{Q}}'_{ik}$ and the points $\mathbf{Q}'_{ik}$ that were originally reconstructed from the depth readings:

$$e_{LRF} = \sum_i \sum_k ||\mathbf{Q}'_{ik} - \widetilde{\mathbf{Q}}'_{ik}||^2$$

The iterative optimization considers the following objective function

$$\min_{\mathsf{T},\mathbf{\Pi}_i} e = e_{LRF} + \kappa\, e_{CAM}, \qquad (21)$$

where e_{CAM} denotes the re-projection error of the plane-to-image homographies, and κ is a weighting parameter that should be adjusted to normalize the variance of camera and LRF residue distributions. The minimization is carried with respect to the extrinsic calibration T and the pose of the inlier planes $\mathbf{\Pi}_i$ expressed in camera coordinates. The camera residue e_{CAM} depends both of the planes $\mathbf{\Pi}_i$ and the camera intrinsics K. We will assume that K is accurately known but, like in [5], this formulation can be potentially used to refine simultaneously the intrinsic and the extrinsic calibration, by considering the independent parameters of K (focal length, skew, aspect ratio and principal point) as variables to be refined.

5 EXPERIMENTS WITH SYNTHETIC DATA

A first set of experiments is conducted in a simulation environment that considers a $0.25°$ resolution LRF and a 1280×960 resolution pin-hole camera. The LRF is assumed to be stationary and the pin-hole camera is randomly placed in a pre-defined region according to a uniform distribution. The camera placement is such that there is always a significant overlap between the fields of view of the two sensors. Given the camera and the LRF, we simulate a set of N checkerboard planes with random poses. Once again the plane placement is such that guarantees intersection with the laser scan plane and a minimum number of grid points visible in the camera. We add gaussian noise to both the image grid points and the laser depth readings. Please note that the pose of the checkerboard plane affects the number of points $\mathbf{Q}'_{ik}$ that are reconstructed by the LRF, and hence the accuracy of the lines $\mathbf{L}'_i$ used for the plane-line registration.

This simulation environment provides input data for performing the extrinsic calibration. The estimations for the relative rotation R and a translation $\mathbf{t}$ are compared with the ground-truth R_{GT} and $\mathbf{t}_{GT}$. The accuracy is typically quantified by the angular magnitude of the residual rotation $\mathsf{R}^\mathsf{T}\mathsf{R}_{GT}$, and by the relative translation

error $||\mathbf{t} - \mathbf{t}_{GT}|| / ||\mathbf{t}_{GT}||$. The results are presented by the Matlab function *boxplot* that shows the two middle quartiles of the distribution (25th to 75th percentiles) as a box with horizontal line at the median. The whisker edges refer to the lowest and highest quartiles, and the crosses show data beyond 1.5 times the interquartile range (outliers in the distribution).

5.1 Extrinsic calibration with minimum data ($N = 3$)

In this experiment the extrinsic calibration is carried using $N = 3$ calibration planes. For each trial, we randomly generate one camera pose and three checkerboard planes. The simulated image points and laser depths are used as calibration input after adding gaussian noise. Since the noise affects both the estimation of planes $\mathbf{\Pi}_i$ and lines $\mathbf{L}'_i$, the result of the plane-line registration is in general different from the correct rigid displacement between the two sensors [1]. Fig. 6 shows the distribution of these errors in 100 independent trials, for increasing amounts of noise in the camera and/or laser. For visualization purposes, we do not plot results without noise, however, we also simulated this situation to perform a sanity check, yelding to maximum errors of 0.0021% for translation and 0.0012 degrees for rotation. The median error is below $10^{-10}\%$ for translation, and below double precision for rotation.

The figure shows compact error distributions with few outliers, which suggests that the calibration algorithm is numerically stable. For low noise levels we can still detect some outliers, that mainly result from the degenerate configuration depicted by situation 1 in figure 4. Input data close to this configuration occur due to generating input planes with bounded orientation variations, so that both sensors can fully detect them. The extrinsic calibration accuracy decreases with increasing amounts of noise, but this degradation is relatively smooth. The rotation estimation seems to be less sensitive to noise than the translation. This is partially explained by the fact that the rotation is computed first and its error propagates to the translation component. In overall terms the results are satisfactory and prove that, if the measurements are not too noisy and the checkerboard orientations are carefully chosen, then the extrinsic camera-LRF calibration can be achieved in practice using only $N = 3$ input planes.

5.2 Comparison with Zhang's method

The simulation framework is now used to compare our algorithm against the calibration method proposed in [5]. For the sake of fairness, Zhang's method is implemented in a hypothesize-and-test framework that is in everything similar to the one described in section 4.1,

---

1. As stated in section 3, the plane-line registration has up to 8 analytical solutions and it is not possible to choose the one corresponding to the extrinsic calibration without further information. In the synthetic experiments we always consider the solution that is closest to the ground-truth

(a) Rotation error

(b) Relative Translation Error

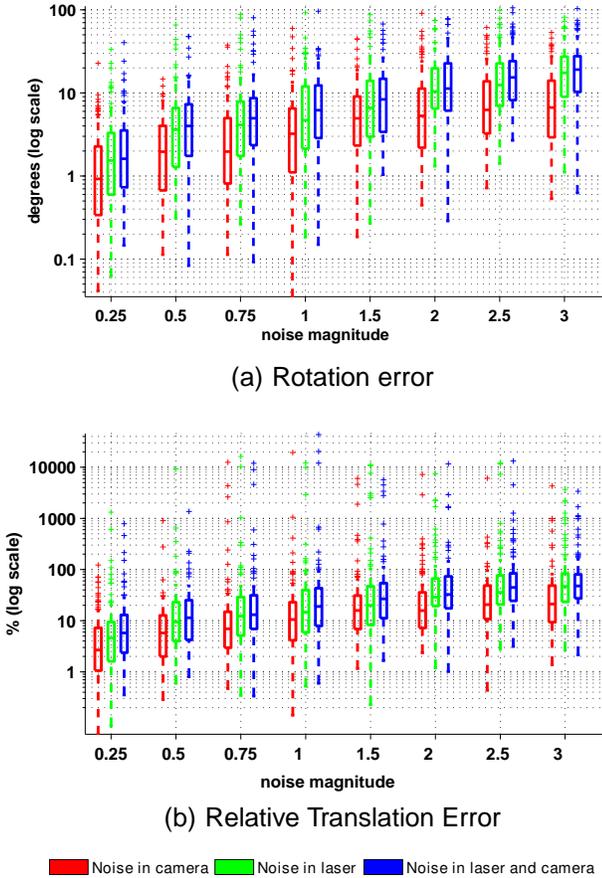| Noise in camera | Noise in laser | Noise in laser and camera |

Fig. 6. Error distribution when the extrinsic calibration is carried using $N = 3$ calibration planes (minimum solution). The labels in the horizontal axis refer to the standard deviation of the added gaussian noise. We consider 1 pixel steps for the camera, and 15 mm steps for the LRF (e.g. a magnitude of 0.25 corresponds to image noise of 0.25 pixels and laser noise of 3.75 mm).

except that it requires 5 plane-line correspondences to generate an hypothesis. The experiment considers a variable number N of calibration planes and, for each case, it runs 100 calibration trials and compares the error distributions. The noise is constant and set to 1 pixel in the camera and 15 mm in the LRF. Fig. 7 shows the results for the two methods before and after running the non-linear optimization discussed in section 4.2.

A careful analysis of the graphics show that our algorithm provides much better initial estimates both in terms of the extrinsic calibration error and residue in the laser. This fact significantly decreases the chances of divergence during the iterative optimization step, specially when the number N of calibration planes is small. It can be observed that, for the case of $N = 5$ and $N = 6$, while our calibration results never diverge significantly from the ground-truth, the final calibration obtained using Zhang's method is often completely erroneous. The stability and final accuracy of the two methods tends

to become similar for a large number of input planes ($N > 8$).

6 EXPERIMENTS WITH REAL DATA

In this experiment we set up a SICK LMS 200 [30] and a camera at fixed positions, and acquire 12 calibration frames by moving a checkerboard pattern in front of the two sensors. The camera intrinsic parameters and the homogeneous coordinates $\mathbf{\Pi}_i$ of the planes are estimated using the intrinsic calibration software described in [12]. Since the re-projection error in the plane-to-image homographies is typically very low (0.2 pixels), there is no advantage in considering the planes $\mathbf{\Pi}_i$ in the final iterative refinement described in section 4.2. Thus, we decided to optimize the cost function of equation 21 only with respect to the relative pose T. Like in the previous experiment, the extrinsic calibration is carried for an increasing number of calibration planes using both our method and the hypothesize-and-test version of Zhang's algorithm. We consider for each N all possible combinations of the 12 frames, which means that the number of trials is

$$\#_N = \frac{12!}{N!\,(12-N)!}$$

In the absence of reliable ground-truth, Fig. 8 shows the distribution of the achieved calibration results. More specifically Fig. 8(a) refers to the angle of the rotation R, Fig. 8(b) concerns the magnitude of the translation vector $\mathbf{t}$, and Fig. 8(c) depicts the LRF residues.

From Fig. 8 it comes that our minimal solution outperforms Zhang's approach when the number of input planes is small. While the former requires 4-to-5 planes for providing accurate estimation results, the latter needs 7 or more planes for achieving a reliable calibration. Fig. 9 confirms that, for the case of $N = 5$, Zhang's extrinsic calibration is in general non plausible. This is justified by the noise in the input data and the existence of frames with few laser readings that are unable to fully constraint the estimation problem. As the number of input planes increases, the hypothesize-and-test procedure discards these frames as outliers, and the output of the two methods converges to the same result. Nevertheless, it is important to remark that this only happens after the bundle adjustment step. While Zhang's initialization is often a coarse estimate of the correct rigid displacement, our closed-form solution is always very close to the global optimum, and the improvements in accuracy achieved by the iterative refinement are somewhat marginal. In addition our approach, being a minimal solution, requires the testing of less hypothesis which is an indisputable advantage in terms of computational complexity. For the case of $N = 12$, an exhaustive search of the solution space requires 220 trials, while the same search with Zhang's method corresponds to 792 tests.

(a) Rotation error

(b) Relative Translation Error

(c) LRF residue

(d) LRF residue (zoom)

▮ Zhang (before optimization)   ▮ Our method (before optimization)   ▮ Zhang (after optimization)   ▮ Our method (after optimization)
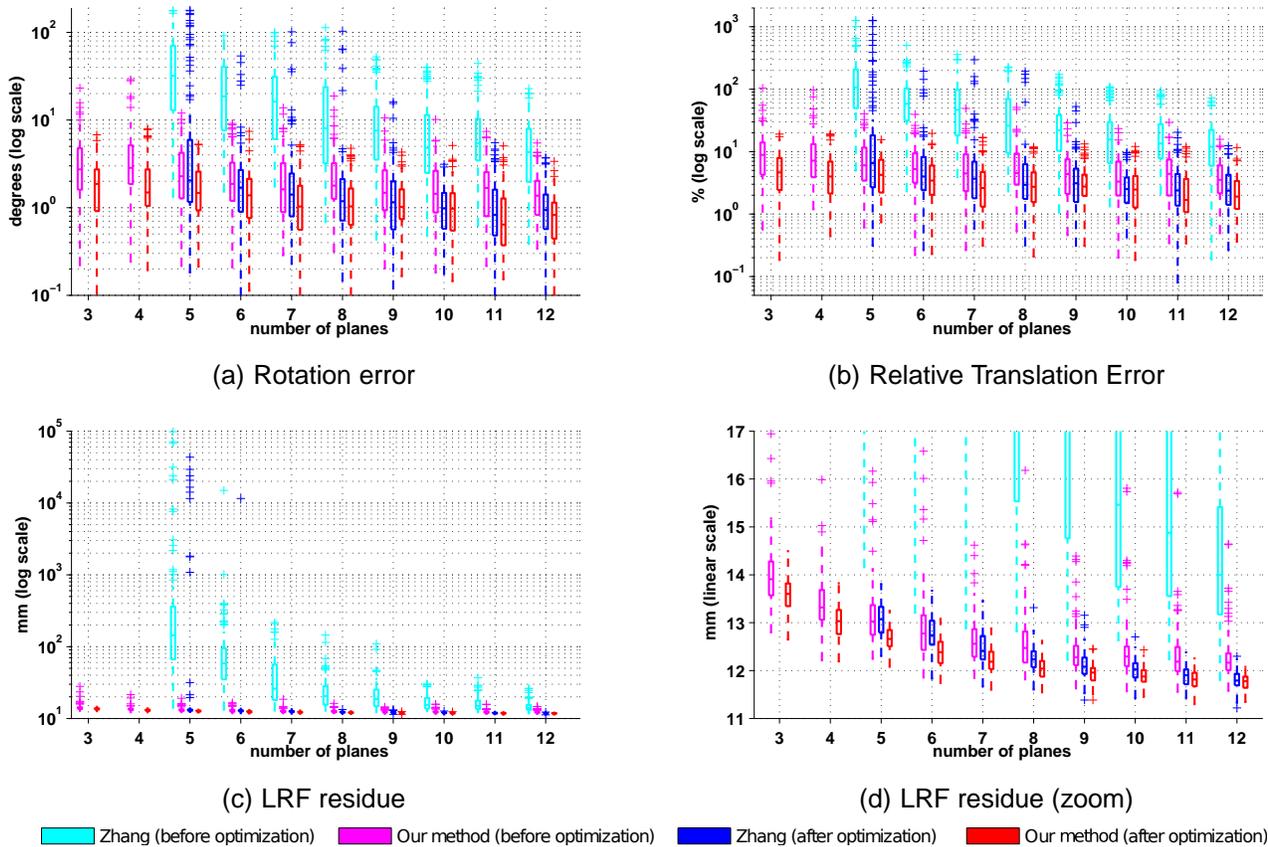
Fig. 7. Calibration using synthetic data. We compare our algorithm against the method presented in [5] when the number N of calibration planes increases. The additive gaussian noise has constant standard deviation of 1 pixel in the camera and 15 mm in the LRF.

7 CONCLUSIONS

This article proposes the first minimal solution for the extrinsic calibration between a camera and an invisible laser-rangefinder. This solution is used as an efficient hypothesis generator in a robust sample-consensus framework. Extensive experiments using both synthetic and real data prove the numerical stability of the minimal formulation, and clearly show that the new calibration algorithm outperforms the state-of-the-art in terms of robustness, accuracy, and required number of input planes.

The core of our calibration method is a new closed-form solution for the registration of 3 planes with 3 coplanar lines in the 3D euclidean space. This is a broad result that is fully characterized and can be useful in other application contexts. As a side result, we provide a more intuitive proof that the problem of relative orientation between two cameras can be formulated as a *p3p* problem [8].

ACKNOWLEDGMENTS
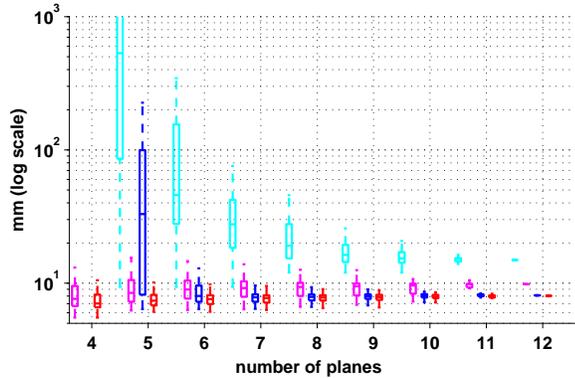
REFERENCES

[1] C. Früh and A. Zakhor, "An automated method for large-scale, ground-based city model acquisition," *International Journal of Computer Vision*, vol. 60, pp. 5–24, 2004.
[2] B. Douillard, D. Fox, and F. Ramos, "Laser and vision based outdoor object mapping," *Robotics: Science and Systems*, 2008.
[3] C. Premebida, O. Ludwig, and U. Nunes, "Lidar and vision-based pedestrian detection system," *Journal of Field Robotics*, vol. 26, pp. 696–711, September 2009.
[4] F. T. Ramos, J. Nieto, and H. F. Durrant-Whyte, "Recognising and modelling landmarks to close loops in outdoor slam." in *IEEE International Conference on Robotics and Automation (ICRA'07)*, 2007, pp. 2036–2041.
[5] Q. Zhang and R. Pless, "Extrinsic calibration of a camera and laser range finder (improves camera calibration)," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'04)*, vol. 3, 2004, pp. 2301 – 2306.
[6] Y. Ma, "An invitation to 3-D vision: from images to geometric models," 2004.
[7] B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon, "Bundle adjustment—a modern synthesis," *Vision Algorithms: Theory and Practice*, pp. 153–177, 1883.
[8] T. Moriya and H. Takeda, "Solving the rotation-estimation problem by using the perspective three-point algorithm," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'00)*, vol. 1, 2000, pp. 766 –773 vol.1.
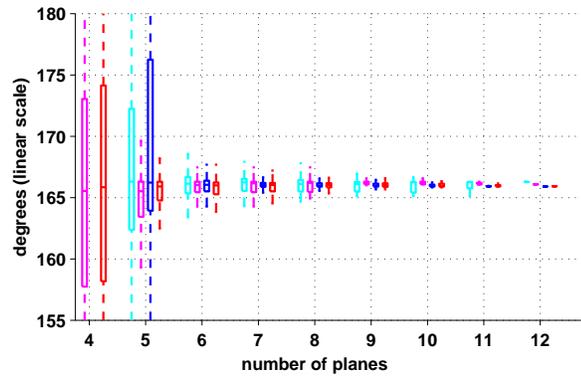
(a) Distribution of the Rotation Angle
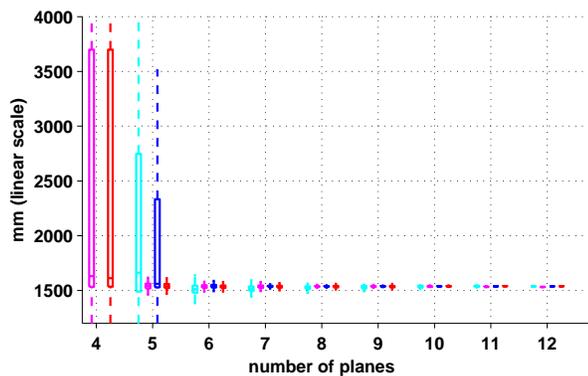
(b) Distribution of the Translation Magnitude

(c) LRF residue

☐ Zhang (before optimization)   ☐ Our method (before optimization)
☐ Zhang (after optimization)   ☐ Our method (after optimization)

Fig. 8. Calibration using real data. The graphics show the distributions of the calibration results obtained using our method and Zhang's algorithm [5]
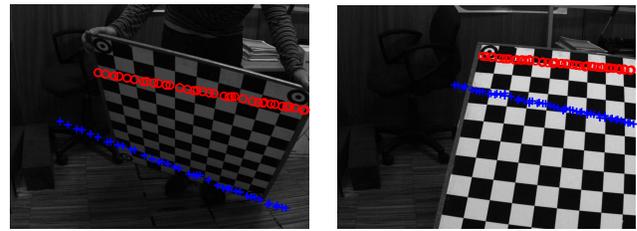

Fig. 9. Projection of the LRF points into the images using the extrinsic calibration results obtained from 5 sampled planes with our method (circles) and Zhang's algorithm (crosses).

[9] R. M. Haralick, C.-N. Lee, K. Ottenberg, and M. Nölle, "Review and analysis of solutions of the three point perspective pose estimation problem," *International Journal of Computer Vision*, vol. 13, pp. 331–356, December 1994.

[10] M. A. Fischler and R. C. Bolles, "Random Sample Consensus - a Paradigm for Model-Fitting with Applications to Image-Analysis and Automated Cartography," *Communications of the Acm*, vol. 24, no. 6, pp. 381–395, 1981.

[11] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," *IEEE International Conference on Computer Vision (ICCV'99)*, vol. 1, pp. 666–673 vol.1, 1999.

[12] J. P. Barreto, J. Roquette, P. Sturm, and F. Fonseca, "Automatic camera calibration applied to medical endoscopy," in *British Machine Vision Conference (BMVC'09)*, 2009.

[13] R. Unnikrishnan and M. Hebert, "Fast extrinsic calibration of a laser rangefinder to a camera," Tech. Rep., 2005.

[14] D. Scaramuzza and A. Harati, "Extrinsic self calibration of a camera and a 3d laser range finder from natural scenes," in *IEEE International Conference on Intelligent Robots and Systems (IROS'07)*, 2007.

[15] C. Gao and J. R. Spletzer, "On-line calibration of multiple lidars on a mobile vehicle platform," in *IEEE International Conference on Robotics and Automation (ICRA'10)*, 2010, pp. 279–284.

[16] G. Li, Y. Liu, L. Dong, X. Cai, and D. Zhou, "An algorithm for extrinsic parameters calibration of a camera and a laser range finder using line features," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, 29 2007-nov. 2 2007, pp. 3854 –3859.

[17] P. Nez, P. D. Jr, R. Rocha, and J. Dias, "Data fusion calibration for a 3d laser range finder and a camera using inertial data," in *European Conference on Mobile Robots (ECMR'09)*, 2009, p. 3136.

[18] H. Aliakbarpour, P. Nuez, J. Prado, K. Khoshhal, and J. Dias, "An efficient algorithm for extrinsic calibration between a 3d laser range finder and a stereo camera for surveillance," in *International Conference on Advanced Robotics (ICAR'09)*, june 2009, pp. 1 –6.

[19] K. D. O. Naroditsky, A. Patterson, "Automatic alignment of a camera with a line scan lidar system," in *IEEE International Conference on Robotics and Automation (ICRA'11)*, 2011.

[20] V. Caglioti, A. Giusti, and D. Migliore, "Mutual calibration of a camera and a laser rangefinder." in *International Conference on Computer Vision Theory and Applications (VISAPP'08)*, 2008, pp. 33–42.

[21] B. K. P. Horn, H. M. Hilden, and S. Negahdaripour, "Closed-form solution of absolute orientation using orthonormal matrices," *Journal of the Optical Society of America A*, vol. 5, no. 7, pp. 1127–1135, Jul 1988.

[22] V. Lepetit, F. Moreno-Noguer, and P. Fua, "Epnp: An accurate o(n) solution to the pnp problem," *International Journal of Computer Vision*, vol. 81, pp. 155–166, February 2009.

[23] C. Olsson, F. Kahl, and M. Oskarsson, "Branch-and-bound methods for euclidean registration problems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 5, pp. 783 –794, May 2009.

[24] S. Ramalingam, Y. Taguchi, T. K. Marks, and O. Tuzel, "P2pi: a minimal solution for registration of 3d points to 3d planes," in *European Conference on Computer Vision (ECCV'10)*. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 436–449.

[25] J. Stolfi, *Oriented Projective Geometry: A Framework for Geometric Computations*. Academic Press, 1991.

[26] J. Grunert, "Das pothenotiscfie problem in erweiterter gestalt nebst ber seine anwendungen in der geodsie," *Gruners Archiv fr Mathematik und Physik*, vol. 1, pp. 238 –248, 1841.

[27] P. F. Sturm and J. P. Barreto, "General imaging geometry for central catadioptric cameras," in *European Conference on Computer Vision (ECCV'08)*, 2008, pp. 609–622.

[28] J. Barreto, J. Santos, P. Menezes, and F. Fonseca, "Ray-based Calibration of Rigid Medical Endoscopes," in *The 8th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras (OMNIVIS'08)*, 2008.

[29] P. H. S. Torr and A. Zisserman, "Mlesac: a new robust estimator with application to estimating image geometry," *Computer Vision and Image Understanding*, vol. 78, pp. 138–156, April 2000.

[30] SICK AG, *LMS200/211/221/291 Laser Measurement Systems*.

Francisco Vasconcelos received the Integrated Master's degree (BSc+MSc) in Electrical and Computers Engineering from the University of Porto, Portugal, in 2009. During 2009 he was a visiting graduate student at the University of Maryland, Baltimore County, USA. Since 2009 he is a computer vision researcher at the Institute for Systems and Robotics – Coimbra. He is currently a PhD student at the University of Coimbra, Portugal. His main research interests lie on geometric computer vision.

João P. Barreto (M'99) received the "Licenciatura" and Ph.D. degrees from the University of Coimbra, Coimbra, Portugal, in 1997 and 2004, respectively. From 2003 to 2004, he was a Postdoctoral Researcher with the University of Pennsylvania, Philadelphia. He has been an Assistant Professor with the University of Coimbra, since 2004, where he is also a Senior Researcher with the Institute for Systems and Robotics. His current research interests include different topics in computer vision, with a special emphasis in geometry problems and applications in robotics and medicine. He is the author of more than 30 peer-reviewed publications. He is also regular reviewer for several conferences and journals, having received 4 *Outstanding Reviewer Awards* in the last few years.

Urbano Nunes (M'95; SM'09) is a Full Professor with the Coimbra University and Vice Director of the Institute for Systems and Robotics, University of Coimbra. Prof. Nunes has been involved with/responsible for several funded projects at both national and international levels in the areas of mobile robotics and intelligent vehicles (such as Cybercars, Cybermove, CyberC3). Prof. Nunes is presently the Vice President for Technical Activities of the IEEE ITS Society, and an Associate Editor of the journals: IEEE Transactions on Intelligent Transportation Systems and IEEE Intelligent Transportation Systems Magazine. He was the General chair for the IEEE ITSC2010 and is appointed as the General Co-chair for IEEE/RSJ IROS 2012.