2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
Congress Center Hamburg
Sept 28 - Oct 2, 2015. Hamburg, Germany

A minimal solution for the Calibration of a 2D Laser-Rangefinder and a Camera based on Scene Corners

Jesus Briales and Javier Gonzalez-Jimenez

***Abstract*— Robots are often equipped with 2D laser-rangefinders (LRFs) and cameras since they complement well to each other. In order to correctly combine the measurements from both sensors, it is required to know their relative pose, that is, to solve their extrinsic calibration. In this paper we present a simple, quick and effective minimal solution for the extrinsic calibration problem. Our approach does not require any on-purpose calibration pattern: it bases on the observation of an orthogonal trihedron, which is profusely found as corners in human-made scenarios. The proposal is validated with synthetic and real experiments, showing better performance than existing alternatives. An implementation of our approach is made available as open-source software[1].**

I. Introduction

The combination of a laser-rangefinder (LRF) and a camera is a common practice in mobile robotics. Some examples are the acquisition of urban models [1], the detection of pedestrians [2], or the construction of semantic maps [3]. Important advances are also being done towards the improvement of visual odometry with additional depth information which can be provided by a LRF as well [4], [5]. In order to effectively exploit the measurements from both types of sensors, a precise estimation of their relative pose, that is, their extrinsic calibration, is required.

The most usual strategy to perform the calibration is to establish some kind of data association between the measurements of both sensors. This, however, is not straightforward owing to the distinctive nature of both sensors: cameras are projective devices which measure angles, whereas 2D LRFs provide range measurements in a single plane. Hence, their calibration is usually posed as a registration problem between different geometric entities coming from the sensors' measurements. On the other hand, to cope with the noise present in the sensors, large sets of data are usually required and the calibration is solved through optimization of some meaningful cost function. Because of the non linear nature of the problem, a variety of methods have been proposed which address such optimization in an iterative manner [6]–[9]. Thus, a good initialization is a key factor for the optimization step in order to assure convergence to the best solution and avoid getting stuck in local minima. Some works propose getting an initial estimate by visual inspection but this, besides limiting accuracy, can be a non trivial task for general users.

The authors are with the Mapir Group —http://mapir.isa.uma.es —of the Department of Systems Engineering and Automation, University of Malaga, Spain. E-mail: jesusbriales@gmail.com

This work has been funded by the Spanish Government under project DPI1011-25483 and the EU FP7 Programme under the GiraffPlus Project.

[1]Code available at http://mapir.isa.uma.es/jbriales

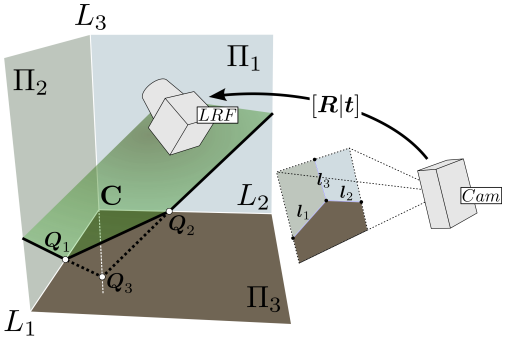


Fig. 1. From the observation of a trihedron structure formed by three orthogonal planes $\{^{\mathcal{T}}\Pi_1, ^{\mathcal{T}}\Pi_2, ^{\mathcal{T}}\Pi_3\}$ and three orthogonal lines $\{L_1, L_2, L_3\}$, three points $\mathbf{Q}_k$ in the 2D LRF plane and three lines $\boldsymbol{l}_k$ in the camera image can be obtained.

Currently, some alternatives have been reported to automatize the obtention of an initial guess for the extrinsic calibration [7], [10]–[12]. However, the potential degeneracies in the underlying registration problem make grabbing the calibration data a key factor in the performance of the method: The stability of the results can be strongly affected by the configuration of recorded data. As a consequence, when the data is adquired in an arbitrary manner, existent methods tend to perform poorly and the chances of getting a wrong estimate increase drastically. In fact, although some guidance may be given to the users on the procedure to follow, the overall process still constitutes an acute challenge to those fresh users, or for end-users from dramatically different fields, who are not aware of the underlying geometrical problem which the calibration relies on.

This paper presents a minimal solution for the described extrinsic calibration that estimates the rotation between the camera and the LRF from *one* single observation of an orthogonal trihedron. The translation, on the other hand, can be fully recovered from the additional observation of a single line (intersection of two scene planes). The only auxiliary element needed, an orthogonal trihedron, can be easily found in any structured scene, e.g. buildings, as the intersection of three perpendicular planes (Fig. 1). This minimal solution is more user-friendly than previous procedures since:

- Fewer observations (only two) are required.
- The results obtained are more robust and stable.
- There is no need to do measurements on the calibration pattern as for checkerboard-based methods.

Experiments conducted using both synthetic and real data support our claim that the proposed calibration method outperforms other existing approaches in terms of reliability

and robustness.

Software and videos derived from this work are available online at: `http://mapir.isa.uma.es/jbriales` . The code includes the minimal solution as well as a GUI for conducting the calibration process.

II. Related Work

Several approaches have been proposed to date which allow us to solve the extrinsic calibration of a camera and a 2D LRF in a closed form. Most of them have based on the method proposed by Zhang and Pless, which in turn relies on the simultaneous observation by both sensors of a checkerboard pattern. In their seminar paper [7], they presented a direct linear solution which requires at least 5 observations. However, this direct solution does not enforce the rotation to fulfill the orthonormality constraint, often leading to a poor initialization for the iterative estimation.

The solution by Zhang and Pless is not minimal, since the extrinsic parameters can be estimated from only three correspondences in generic configuration. This fact is pointed out by Vasconcelos *et al.* in [11], where the authors exploit the existent duality between points and planes in the projective space $\mathbb{P}^3$ to reformulate the geometrical registration problem as a Perspective-3-Point (P3P) problem whose solution has been thoroughly studied over the years [13].

Zhou [12] proposes another minimal solution to the Zhang and Pless' problem which directly exploits the algebraic structure of the corresponding polynomial system in terms of rotation matrix elements, yielding more numerically stable results than Vasconcelos *et al.*

Another distinct minimal solution is proposed for the extrinsic calibration in [10] based on the laser reflectivity information to establish data association between image lines and LRF points. Six of these correspondences compound the minimal problem. Their algorithm, as Zhou's, employs an algebraic approach to solve a set of polynomial equations. However, the reflectivity information is not always available for the LRF (depending on the device), which reduces the generality of this approach.

Other approaches exist already which exploit an structured ambient to perform the extrinsic calibration of different sensors. In [14], scene planes are used to generate geometric constraints which allow for the calibration of non-overlapping camera and LRF devices. In our previous work [15] the observations of an orthogonal trihedron were used to produce constraints and optimize the extrinsic calibration of camera and LRF in an iterative approach. Other methods have also benefitted of structured scenes to set geometric constraints for the calibration of different sensor rigs, e.g. for several LRFs [16].

To the best of our knowledge, ours is the first approach which proposes a *minimal solution* directly exploiting the scene structure. Namely, this constitutes the counterpart minimal solution to the approach in [15] based on scene corners, so both could be readily fused to improve the overall performance.

III. Trihedron Registration

The observation of an orthogonal trihedron from a camera or a LRF has some unique traits unseen in other usual calibration patterns. Next, we present the data which can be extracted from the observation of the trihedron, either by a camera or by a LRF device.

First of all, let us introduce a mathematical characterization of the orthogonal trihedron that will be fully exploited in next sections. The orthogonal trihedron $\boldsymbol{\tau}$ can be mathematically defined as a structure formed by three perpendicular planes, denoted by $\{{}^{\boldsymbol{\tau}}\Pi_1, {}^{\boldsymbol{\tau}}\Pi_2, {}^{\boldsymbol{\tau}}\Pi_3\}$, which intersect in a single vertex $\mathbf{C}$ (see Fig. 1). On the other hand, the trihedron can be alternatively seen as three orthogonal lines, $\{ L_1, L_2, L_3 \}$, also meeting in the vertex $\mathbf{C}$. Both characterizations are tightly connected since each line is produced by the intersection of two planes and each plane is defined by a pair of lines. It is this duality that allows us to exploit the observation of the trihedron by both sensors, since the camera can easily observe 3D lines but not planes and, on the contrary, the LRF can sample planes but not lines.

A. Camera Observation

Each line L_k in the trihedron projects to the camera image as $\boldsymbol{l}_k$. Let us define the *interpretation plane* of the line $\boldsymbol{l}$ as the plane which passes through the camera origin and the image line $\boldsymbol{l}$ (Fig. 2). Due to the projective nature of

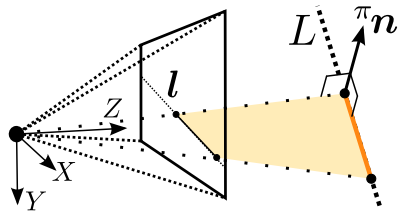


Fig. 2. For a calibrated camera, the image line $\boldsymbol{l}$ defines, together with camera center, the *interpretation plane* in which L must lie.

the camera the original 3D line L_k can not be recovered from $\boldsymbol{l}_k$, but it must lie in its corresponding *interpretation plane*. Let $\boldsymbol{l}$ stand for the homogeneous representation in $\mathbb{P}^2$ of the corresponding 3D line. We assume the camera has been previously calibrated, so that $\boldsymbol{K}$ stands for the intrinsic calibration matrix of the pinhole camera model. Then, the homogeneous $\mathbb{P}^3$ representation of the interpretation plane for $\boldsymbol{l}$ is

$$ {}^{\boldsymbol{\pi}}\Pi = \begin{bmatrix} {}^{\boldsymbol{\pi}}\boldsymbol{n} \\ 0 \end{bmatrix}, \quad {}^{\boldsymbol{\pi}}\boldsymbol{n} = \frac{\boldsymbol{K}^\top \boldsymbol{l}}{\|\boldsymbol{K}^\top \boldsymbol{l}\|} \tag{1} $$

In a similar fashion, the vertex image $\boldsymbol{c}$ back-projects to a ray whose unitary direction $\hat{\boldsymbol{c}}$ is defined by

$$ \hat{\boldsymbol{c}} = \frac{\boldsymbol{K}^{-1}\tilde{\boldsymbol{c}}}{\|\boldsymbol{K}^{-1}\tilde{\boldsymbol{c}}\|}, \quad \tilde{\boldsymbol{c}} = \begin{bmatrix} \boldsymbol{c} \\ 1 \end{bmatrix} \tag{2} $$

Overall, the orthogonal trihedron projects onto the image as a set of three intersecting half-lines $\{\boldsymbol{l}_k\}_{k=1}^3$ (see Fig. 4(a)) which can be manually hinted and then accurately adjusted with the approach proposed in [15]. Let us define $\boldsymbol{N}_{\boldsymbol{\tau}}$ as

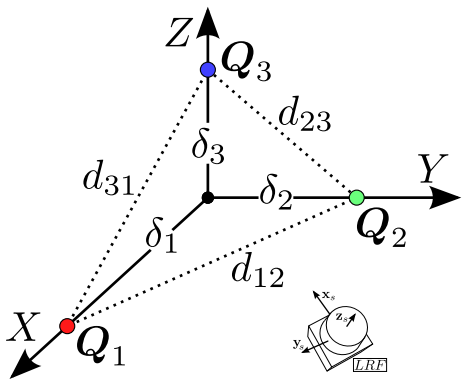


Fig. 3. Solving the LRF pose from the intersection of its plane with an orthogonal trihedron is equivalent to a P3P problem with orthogonal angles.

the orientation of the 3D lines in the orthogonal trihedron with respect to (wrt) the camera coordinate system. The orientation $\boldsymbol{N}_{\boldsymbol{\tau}}$ can be solved from the lines $\{\boldsymbol{l}_k\}_{k=1}^3$ up to a mirror transformation [17]. If we set an orientation for each line, so that $\det(\boldsymbol{N}_{\boldsymbol{\tau}}) = +1$, the solution is unique.

As a conclusion, one single image of the orthogonal trihedron permits us to recover its pose wrt the camera up to a scale factor λ in the translation, due to the projective nature of the camera:

$$ {}^{c}\boldsymbol{R}_{\boldsymbol{\tau}} = \boldsymbol{N}_{\boldsymbol{\tau}}, \quad {}^{c}\boldsymbol{t}_{\boldsymbol{\tau}} = \lambda\, \hat{\boldsymbol{c}} \tag{3} $$

B. LRF Observation

Each plane ${}^{\tau}\Pi_k$ in the trihedron is sampled by the LRF as a set of collinear points $\{\boldsymbol{p}_k^i\}_{i=1}^{N_k}$ which can be used to fit the 2D line $\boldsymbol{m}_k$ corresponding to the intersection of the LRF plane with ${}^{\tau}\Pi_k$. As can be noted from Figs. 1 and 4(b), the intersection of each pair of lines $\boldsymbol{m}_i$ and $\boldsymbol{m}_j$ provides the 2D point $\boldsymbol{q}_k$ corresponding to the intersection of the LRF plane with L_k. As a result, when all the three planes of the trihedron are sampled by the LRF, all the intersection points $\{\boldsymbol{q}_k\}_{k=1}^3$ can be computed.

Let us consider, for the rest of this section, the system defined for the trihedron as the reference coordinate system. The points $\mathbf{Q}_k$, 3D counterparts of the 2D intersection points $\boldsymbol{q}_k$, have a simple parameterization wrt the trihedron basis given by

$$ \mathbf{Q}_k = \delta_k \mathbf{e}_k, \tag{4} $$

with $\mathbf{e}_k$ the k-th canonical vector and δ_k an unknown signed distance. On the other hand, the distance d_{ij} between each pair of points $\mathbf{Q}_i$ and $\mathbf{Q}_j$ can be recovered from the LRF measurements as

$$ d_{ij} = \left\| \boldsymbol{q}_i - \boldsymbol{q}_j \right\|. \tag{5} $$

Hence, the information available defines a classical Perspective-3-Points (P3P) problem (see Fig. 3): Given the distances between points $\mathbf{Q}_1$, $\mathbf{Q}_2$ and $\mathbf{Q}_3$ and the 3D rays in which these points lie, determine the unknown depths δ_1, δ_2 and δ_3. This problem has been thoroughly studied in the computer vision literature [13]. Nevertheless, the orthogonal nature of the rays in this case greatly simplifies the complexity of the problem, which reduces to

$$ \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \delta_1^2 \\ \delta_2^2 \\ \delta_3^2 \end{bmatrix} = \begin{bmatrix} d_{23}^2 \\ d_{31}^2 \\ d_{12}^2 \end{bmatrix} \tag{6} $$

Only $|\delta_k|$ can be recovered from the information available, so eight possible solutions exist because of the different permutations of signs. However, the user can easily set the correct sign for each δ_k through visual inspection of the scene and the LRF pose. The set then reduces to one feasible solution $\{\delta_1, \delta_2, \delta_3\}$.

Let us assume, without loss of generality, that the LRF plane is defined by $z = 0$ in LRF coordinate system, with X direction pointing forwards. The relation between trihedron points $\mathbf{Q}$ and the LRF 2D points then reduces to

$$ \mathbf{Q} = \begin{bmatrix} \boldsymbol{r}_x & \boldsymbol{r}_y & {}^{\tau}\boldsymbol{t}_s \end{bmatrix} \begin{bmatrix} \boldsymbol{q} \\ 1 \end{bmatrix} \tag{7} $$

Once both $\{\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3\}$ and $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3\}$ have been recovered, the components of the LRF pose wrt trihedron are easily determined as

$$ \begin{bmatrix} \boldsymbol{r}_x & \boldsymbol{r}_y & {}^{\tau}\boldsymbol{t}_s \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_1 \, \mathbf{Q}_2 \, \mathbf{Q}_3 \end{bmatrix} \begin{bmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 & \boldsymbol{q}_3 \\ 1 & 1 & 1 \end{bmatrix}^{-1} \tag{8} $$

$$ {}^{\tau}\boldsymbol{R}_s = \begin{bmatrix} \boldsymbol{r}_x & \boldsymbol{r}_y & \boldsymbol{r}_x \times \boldsymbol{r}_y \end{bmatrix} $$

In conclusion, one single observation of all the trihedron planes allows us to fully recover the pose of the LRF wrt the trihedron.

IV. The Calibration Problem

The extrinsic calibration of a camera and a LRF consists in finding the 3D rigid transformation between the LRF coordinate system and the camera coordinate system. Let ${}^{c}\boldsymbol{p}$ and ${}^{s}\boldsymbol{p}$ be the coordinates of a 3D point as seen from the camera and the LRF references, respectively. Once extrinsic calibration is performed the relation

$$ {}^{c}\boldsymbol{p} = \boldsymbol{R}\, {}^{s}\boldsymbol{p} + \boldsymbol{t} \tag{9} $$

stands, where $\boldsymbol{R}$ denotes the orientation of the LRF relative to the camera, and $\boldsymbol{t}$ the translation vector between both sensors, with the z component of the scan points ${}^{s}\boldsymbol{p}$ conventionally set to zero.

As shown in Section III, the unique characteristics of a single trihedron observation from camera and LRF devices allows us to almost fully recover the relative pose between the trihedron and the devices. Consequently, solving the extrinsic calibration of the LRF and the camera once the registration of the trihedron wrt to each device is performed becomes as simple as composing the poses:

$$ \boldsymbol{R} = {}^{c}\boldsymbol{R}_{\boldsymbol{\tau}} \cdot {}^{\tau}\boldsymbol{R}_s, \quad \boldsymbol{t} = \lambda \hat{\boldsymbol{c}} + {}^{c}\boldsymbol{R}_{\boldsymbol{\tau}} \cdot {}^{\tau}\boldsymbol{t}_s \tag{10} $$

The expressions above is based on the results (3) and (8) and completely define the rotation between both sensors. However, an unknown scale factor λ remains for ${}^{c}\boldsymbol{t}_{\boldsymbol{\tau}} = \lambda \hat{\boldsymbol{c}}$. This can be solved by the observation of an additional *line* L_*. This provides an additional point $\boldsymbol{q}_*$ in the LRF plane

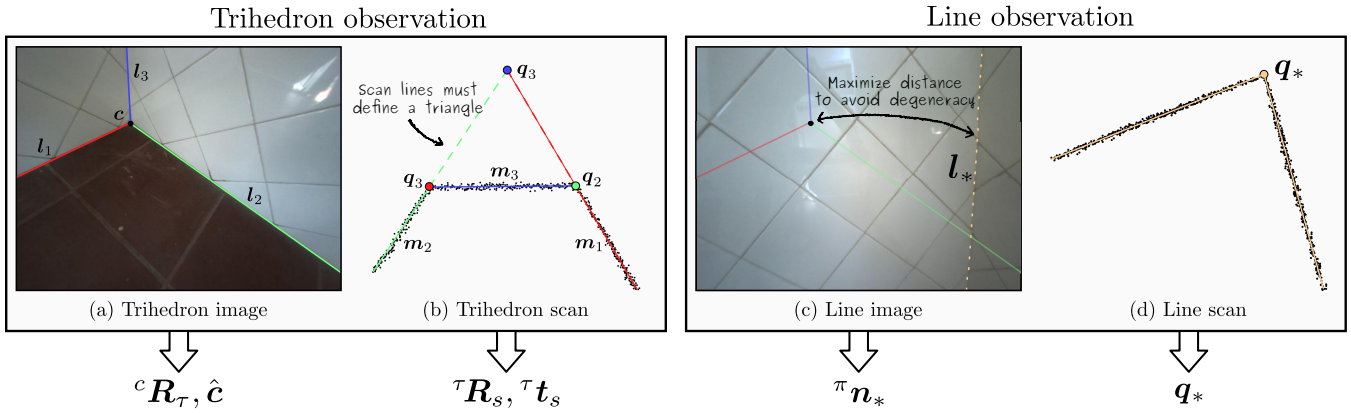


Fig. 4. Flow chart of the proposed minimal solution: The camera pose (up to scale) and the LRF pose wrt trihedron are recovered from a *single* trihedron observation (a,b) and enables us to fully recover the rotation part of the calibration. The unknown scale factor is solved with an additional line observation (c,d) to get the translation component of the calibration. See the supplementary video for best visualization of the algorithm pipeline and its real use.

and a new interpretation plane ${}^{\pi}\Pi_{*}$ for the camera (see Figs. 4(c) and 4(d)). For a point lying in a line that is in turn contained in a plane, the point also belongs to the plane:

$$\left.\begin{array}{c}\boldsymbol{q}_{*} \in L_{*} \\ L_{*} \in {}^{\pi}\Pi_{*}\end{array}\right\} \Rightarrow \boldsymbol{q}_{*} \in {}^{\pi}\Pi_{*} \tag{11}$$

This defines the *point-to-plane* constraint which, after transformation of $\boldsymbol{q}_{*}$ into camera coordinates (9), reads

$$\left({}^{\pi}\boldsymbol{n}_{*}\right)^{\top}\left(\boldsymbol{R}\,\boldsymbol{q}_{*}+\boldsymbol{t}\right)=0. \tag{12}$$

The substitution of (10) into (12) allows us to solve the unknown scale parameter:

$$\lambda=-\frac{\left({}^{\pi}\boldsymbol{n}_{*}\right)^{\top}\left(\boldsymbol{R}\,\boldsymbol{q}_{*}+{}^{c}\boldsymbol{R}_{\tau}\,{}^{\tau}\boldsymbol{t}_{s}\right)}{\left({}^{\pi}\boldsymbol{n}_{*}\right)^{\top}\hat{\boldsymbol{c}}}. \tag{13}$$

The solution of λ becomes degenerate when the denominator in (13) is zero. This situation, however, can be easily avoided by taking the additional line far from the previously observed trihedron vertex, as pointed in Fig. 4(c).

V. Experimental Results

In this section the proposed minimal solution is validated through a number of experiments with both synthetic and real data. Our results are compared to those obtained by the state-of-the-art alternative minimal solution by Vasconcelos *et al.* [11] which bases on the use of a checkerboard pattern.

A. Simulation environment

In the first set of experiments a simulation environment is used to assess the proposed solution in terms of numerical precision and robustness. A rig formed by a LRF and a camera is randomly generated with extrinsic calibration in the range of $\pm 45^{\text{o}}$ and ± 50 cm for the rotation and translation respectively. The intrinsic parameters for the simulated sensors are taken from the Hokuyo UTM-30LX LRF and the Point Grey Bumblebee®2 stereo camera employed in the subsequent real experiments.

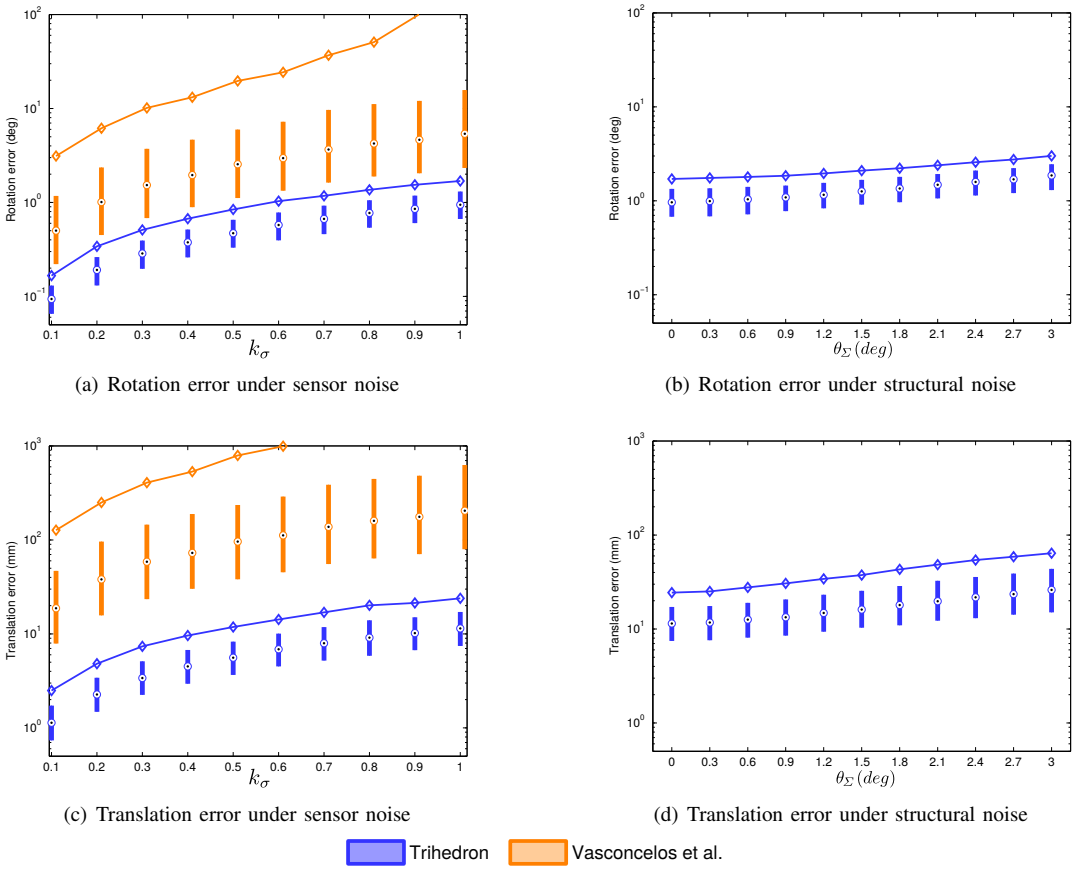
Two different targets are used to simulate the observations necessary for our method and [11]: a flat orthogonal trihedron formed by three $L \times L$ squares, and a $L \times L$ checkerboard with 8×8 squares, respectively, with $L = 1.5$ m. Notice that, whereas a similar size has been taken for both targets, the trihedron structures present in most structured scenes have bigger dimensions, thus providing extra robustness in the extraction of scan features.

The rig is randomly placed in a predefined region according to a uniform distribution, and the poses are checked to guarantee the correctness of all the simulated observations. The observations are affected by variable levels of unbiased and uncorrelated Gaussian noise, σ_{S} and σ_{C}, for both scan range measurements and image points, respectively. In order to assess the accuracy of the results, the common error metrics

$$e_{\boldsymbol{R}}=2\arcsin\left(\frac{\left\|\boldsymbol{R}-\hat{\boldsymbol{R}}\right\|_{\mathfrak{F}}}{2\sqrt{2}}\right), \quad e_{\boldsymbol{t}}=\left\|\boldsymbol{t}-\hat{\boldsymbol{t}}\right\|,$$

are employed, where $[\boldsymbol{R}|\boldsymbol{t}]$ is the estimated transformation and $[\hat{\boldsymbol{R}}|\hat{\boldsymbol{t}}]$ is the ground truth. The metric $e_{\boldsymbol{R}}$ stands for the lowest angular distance, in radians, between two rotations in $SO(3)$. Translation metric $e_{\boldsymbol{t}}$ is the Euclidean distance between two $\mathbb{R}^3$ vectors.

The level of noise in the measurements has been modeled by a variable factor of proportionality k_{σ} which multiplies the most usual standard deviation of measurements for each sensor, i.e. $\sigma_{C} = 1$ pixel for image points and $\sigma_{S} = 30$ mm for range measurements in the LRF. A statistical analysis of the errors is shown in Figs. 5(a) and 5(c) for an increasing level of noise k_{σ}. We observe that our solution clearly outperforms that of Vasconcelos *et al.* in precision by almost an order of magnitude in the case of rotation and more than a decade for translation. Our solution, however, introduces the assumption that the trihedron observed is orthogonal, thus a new source of error appears whose effect is evaluated too. For that purpose, each trihedron line is perturbed by certain angle in a random direction in its corresponding orthogonal plane. In order to quantify the introduced perturbation we propose the following intuitive metric: For a given set of non-orthogonal 3D directions stacked as columns in $\boldsymbol{V}$, compute its closest rotation matrix $\bar{\boldsymbol{V}}$ by Singular Value Decomposition, then sum the angular distances for each direction in the trihedron wrt the corresponding rotation



(a) Rotation error under sensor noise

(b) Rotation error under structural noise

(c) Translation error under sensor noise

(d) Translation error under structural noise

Fig. 5. Simulated calibration results (each step comprises 5000 samples). The trend line of 90th percentile is shown as a measure of robustness. (a)(c) Comparison with [11] for increasing level of sensor noise. (b)(d) Assessment of performance when orthogonal assumption does not hold, for increasing levels of maximum perturbation angle.

direction to get the perturbation measure

$$\theta_\Sigma = \sum_{k=1}^{3} \arccos(\boldsymbol{V}_k^\top \bar{\boldsymbol{V}}_k)$$

The effect of this perturbation on both the rotation and the translation is shown in Figs. 5(b) and 5(d) for an increasing maximum perturbation angle θ_Σ which is equally divided among three directions, being the sensor noise fixed to their base values ($k_\sigma = 1$). The simulation results evidence that even if the orthogonality condition is not completely fulfilled the results obtained outperform those of Vasconcelos *et al.*

If the minimal solution is seen as a means of initialization for iterative approaches, robustness is a highly desirable characteristic: A lower precision can be saved by a few additional steps in the iterative stage, whereas a initial state based on a bad estimate could drive to false minima. For this reason, even though outliers are not plotted for the sake of better visualization, the 90th percentile is also plotted in the boxplots in Fig. 5 as a measure of robustness. The 90th percentile of the commited error shows that the improvement of our approach in terms of robustness is even greater than it is for precision (please note the logarithmic scale of the boxplot), with a difference of more than two orders of magnitude under usual noise conditions. This improvement keeps even for usual disturbances in the orthogonal assumption.

B. Real data

The rig employed in our experiments is composed of a Hokuyo UTM-30LX LRF and a Point Grey Bumblebee®2 stereo camera (see Fig. 6). Using these sensors, an overall

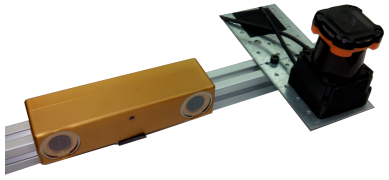


Fig. 6. Rig employed in the real simulations formed by a Point Grey Bumblebee®2 camera (left) and a Hokuyo UTM-30LX LRF (right).

of 500 trihedron observations were taken from various poses and in different scenes. On the other hand, 16 checkerboard observations were taken from different poses too for testing the approach by Vasconcelos *et al.* [11].

Due to the lack of precise ground truth data for the extrinsic calibration of the sensor rig, various approaches were exploited to prove the robustness and accuracy of our minimal solution. First, the extrinsic calibration of the left sensor in the stereo camera and the LRF was performed applying our minimal solution to each of the 500 trihedron observations gathered together with an additional line observation. On the other hand, 500 independent calibrations were performed

with the method in [11] using minimal combinations of 3 checkerboard observations plus an additional one for solution disambiguation. Then, a visual validation on the robustness of both methods was performed in which the LRF points lying in the surface of a white planar board were projected to image using every available calibration result. The points projected with our calibration results fit well within the board limits (Fig. 7). Furthermore, the greater constancy and robustness of our solutions (green) can be checked against those of Vasconcelos *et al.* for which reprojection on the image varies considerably producing many clearly invalid reprojections. A second approach to measure the robustness

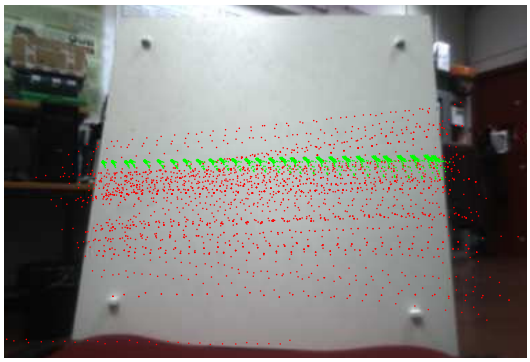


Fig. 7. Projection of the LRF points on a white planar board into its image using several calibration results obtained from minimal sets of data with our method (green) and [11] (red). Best seen in color.

of the methods relies on the calibration of the stereo camera, which is provided by the manufacturer and can be considered as a reliable ground truth. Then, similarly to [15] and [10], the relative transformation between left and right cameras is computed indirectly through the extrinsic calibration wrt the LRF of each camera. The stereo evaluation for the set of 500 calibrations, depicted in Fig. 8, supports the conclusions from the visual test: our solution shows more robustness, with typically lower loop closure errors and the error associated to outliers remarkably lower than for the solution provided by [11].

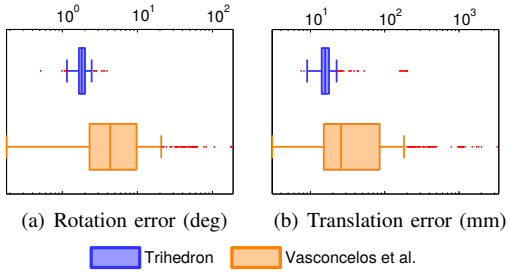


Fig. 8. Statistic of the stereo loop closure error from 500 indivuals calibrations. Please be aware of the logarithmic scale in the error axis.

VI. Conclusions

In this paper, a new minimal solution has been presented for the extrinsic calibration of a rig formed by a 2D LRF and a camera. It relies on the observation of an orthogonal trihedron, which is typically encountered as scene corners in buildings and human-made structures. Namely, our approach treats the resolution of each sensor pose wrt the trihedron as an independent registration problem, and then fuses this information to solve the extrinsic calibration of the rig. Our calibration technique reduces the number of necessary observations and, at the same time, avoids the need of a calibration pattern built on purpose for the task. The method has been extensively tested in simulation and in real case experiments, where it achieves both increased robustness and repeatability wrt the state-of-the-art minimal solution.

References

[1] J.-L. Blanco, F.-A. Moreno, and J. Gonzalez, "A collection of outdoor robotic datasets with centimeter-accuracy ground truth," *Autonomous Robots*, vol. 27, no. 4, pp. 327–351, 2009.

[2] C. Premebida and U. J. C. Nunes, "Fusing LIDAR, camera and semantic information: A context-based approach for pedestrian detection," *The International Journal of Robotics Research*, p. 0278364912470012, 2013.

[3] B. Douillard, D. Fox, F. Ramos, and H. Durrant-Whyte, "Classification and semantic mapping of urban environments," *The international journal of robotics research*, vol. 30, no. 1, pp. 5–32, 2011.

[4] J. Zhang, M. Kaess, and S. Singh, "Real-time depth enhanced monocular odometry," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pp. 4973–4980, Sept. 2014.

[5] C. Bodensteiner, W. Hubner, K. Jungling, P. Solbrig, and M. Arens, "Monocular Camera Trajectory Optimization using LiDAR data," *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, pp. 2018–2025, Nov. 2011.

[6] S. Wasielewski and O. Strauss, "Calibration of a multi-sensor system laser rangefinder/camera," in *Intelligent Vehicles '95 Symposium., Proceedings of the*, pp. 472–477, Sept. 1995.

[7] Q. Zhang and R. Pless, "Extrinsic calibration of a camera and laser range finder (improves camera calibration)," in *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, vol. 3, pp. 2301–2306 vol.3, Sept. 2004.

[8] K. Kwak, D. F. Huber, H. Badino, and T. Kanade, "Extrinsic calibration of a single line scanning lidar and a camera," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pp. 3283–3289, Sept. 2011.

[9] P. Tulsuk, P. Srestasathiern, M. Ruchanurucks, T. Phatrapornnant, and H. Nagahashi, "A novel method for extrinsic parameters estimation between a single-line scan LiDAR and a camera," in *Intelligent Vehicles Symposium Proceedings, 2014 IEEE*, pp. 781–786, 2014.

[10] O. Naroditsky, A. Patterson, and K. Daniilidis, "Automatic alignment of a camera with a line scan LIDAR system," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pp. 3429–3434, May 2011.

[11] F. Vasconcelos, J. P. Barreto, and U. Nunes, "A Minimal Solution for the Extrinsic Calibration of a Camera and a Laser-Rangefinder," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, pp. 2097–2107, Nov. 2012.

[12] L. Zhou, "A New Minimal Solution for the Extrinsic Calibration of a 2D LIDAR and a Camera Using Three Plane-Line Correspondences," *Sensors Journal, IEEE*, vol. 14, pp. 442–454, Feb. 2014.

[13] M. Haralick, Robert M. and Lee, Chung-Nan and Ottenberg, Karsten and Nölle, C.-N. C. Lee, K. Ottenberg, M. Nölle, and B. Haralick, "Review and analysis of solutions of the three point perspective pose estimation problem," *International journal of computer vision*, vol. 356, no. 1994, pp. 331–356, 1994.

[14] Y. Bok, D.-G. Choi, P. Vasseur, and I. S. Kweon, "Extrinsic calibration of non-overlapping camera-laser system using structured environment," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pp. 436–443, Sept. 2014.

[15] R. Gomez-Ojeda, J. Briales, E. Fernandez-Moral, , and J. Gonzalez-Jimenez, "Extrinsic Calibration of a 2D Laser-Rangefinder and a Camera based on Scene Corners," in *International Conference on Robotics and Automation (ICRA 2015)*, IEEE, 2015.

[16] E. Fernández-Moral, V. Arévalo, and J. González-Jiménez, "Extrinsic calibration of a set of 2D laser rangefinders," in *International Conference on Robotics and Automation (ICRA 2015)*, IEEE, 2015.

[17] K. Kanatani, "Constraints on length and angle," *Computer Vision, Graphics, and Image Processing*, pp. 28–42, 1988.