ACCEPTED VERSION

Alireza Khosravian, Tat-Jun Chin, Ian Reid

A branch-and-bound algorithm for checkerboard extraction in camera-laser calibration

Published version at: http://dx.doi.org/10.1109/ICRA.2017.7989767

http://hdl.handle.net/2440/112401

A Branch-and-Bound Algorithm for Checkerboard Extraction in Camera-Laser Calibration

Alireza Khosravian, Tat-Jun Chin, Ian Reid

Abstract— We address the problem of camera-to-laser-scanner calibration using a checkerboard and multiple image-laser scan pairs. Distinguishing which laser points measure the checkerboard and which lie on the background is essential to any such system. We formulate the checkerboard extraction as a combinatorial optimization problem with a clear cut objective function. We propose a branch-and-bound technique that deterministically and globally optimizes the objective. Unlike what is available in the literature, the proposed method is not heuristic and does not require assumptions such as constraints on the background or relying on discontinuity of the range measurements to partition the data into line segments. The proposed approach is generic and can be applied to both 3D or 2D laser scanners as well as the cases where multiple checkerboards are present. We demonstrate the effectiveness of the proposed approach by providing numerical simulations as well as experimental results.

I. Introduction

Many robotics applications rely on a camera and a laser scanner, which are rigidly attached together, as the main sensors for performing localization, navigation, and mapping the surrounding environment. Effective fusion of laser measurements with camera images requires knowledge of the extrinsic transformation between the two sensors.

The process of calibrating either a 2D or 3D laser scanner to a camera by computing the rigid transformation between their reference frames is known as *camera-laser calibration*. Various methods have been developed for this [1]–[12], including free toolboxes (e.g. [8] for 2D lasers and [7], [11] for 3D lasers). Although a few checkerboard free calibration methods have been developed in the literature (see e.g. [2], [4]–[6]), still a very popular approach is to place one or multiple known targets (e.g. planar checkerboards) in the scene such that they are observed in both the image and the laser scans (see e.g. [1], [3], [7]–[11], [13]–[15]). We take the later approach in this paper. In this approach, one or multiple sets of laser-image pairs are collected. The checkerboard corners associated with each image are identified and the checkerboard plane is obtained in the camera coordinate frame. Then, the laser scan points that hit the checkerboard plane in each image are extracted and used to solve a nonlinear optimization that estimates the extrinsic transformation between the camera and laser scanner.

An important part of the process is *checkerboard extraction*, in which we must determine which laser scan points

fall on the checkerboard and which lie on the background. Manual extraction of those inliers is a time consuming and inaccurate process, although some popular toolboxes still rely on this manual extraction (see e.g. [7]). Automatic checkerboard extraction has been poorly studied in the literature and [8] seemed to be the only reference that investigates this problem in details. Other references, only briefly discuss the checkerboard extraction as a block in the pipeline of the camera-laser calibration process [3], [9], [11]. All of these methods are heuristic in nature, imposing assumptions on the environment or the camera-laser setup to identify possible candidate laser points that might correspond to the checkerboard. Those candidates are then passed to a randomised hypothesis-and-test search (e.g. RANSAC [16]) to eliminate the outliers and identify the correspondence of the laser points and the checkerboards. The imposed assumptions depend on whether a 2D or 3D laser scanner is employed and are specific to particular situations. For instance, where 2D laser scanner is employed, [8], [9] and [3, Section 4] rely on range discontinuities to partition the range measurements into line segments that might correspond to a planar checkerboard. Also [8] additionally assumes that the background is stationary in most of the scans and relies on the frequency of occurrence of range measurements. Assumptions on the minimum and maximum distance of the checkerboard to the laser scanner during different scans, have also been used to enhance the checkerboard extraction process [11].

The heuristic methods have demonstrated good performance in some applications [3], [8], [9], [11], [16]. However, imposing assumptions limits the application of these methods to particular scenarios. Assumptions such as stationary background limit the usage of the developed techniques to the controlled environments, preventing their application to open spaces where the background could inevitably change due to moving objects. Also, in some practical scenarios, stationary known targets are used for calibration and the camera-laser pair (usually mounted on a vehicle) is moved around and collects the data. In such situations, the background does inevitably change. Relying on range discontinuities is limited to the cases where the checkerboard is not placed on the walls and there are not many other objects present in the scene that might cause unwanted range discontinuities. Little attention has been given in literature to the design of *systematic* methods for checkerboard extraction based on rigorous theoretical foundations that can perform robustly in all practical situations.

Branch-and-bound (BnB) is a systematic method for discrete and combinatorial optimization. In the context of geo-

metric model fitting, BnB is a popular technique in situations where outliers are present or measurement correspondences are unknown (see e.g. [17]–[20] and the references therein). Building on a rigorous theoretical foundation, this method systematically searches the state space for inliers that optimize an objective function and is guaranteed to find the globally optimal state. A BnB method has recently been developed for the hand-eye calibration problem [21], [22] and demonstrated successful application. To the best of our knowledge, the BnB method has not been applied to the camera-laser calibration problem prior to the current contribution.

In this paper, we propose a BnB technique for checkerboard extraction in the camera-laser calibration problem. Following [1], we assume that checkerboard normals are obtained in the camera coordinate frame by processing the checkerboard corners. We propose an objective function that represents the number of laser points that fall within a box with a small distance to the checkerboard (inliers). A key contribution of the paper is to propose a tight upper bound for the objective function that ensures the BnB method finds the globally optimal inlier laser points. Our method is not heuristic and relies only on the underlying geometry of the camera-laser calibration problem without imposing any assumption or constraints on the environment. Hence it is quite general and performs robustly in all practical situations. It is applicable directly to both 2D and 3D laser scanners, to the case where multiple checkerboards are present in the scene, as well as the case where some of the checkerboards either do not fall or only partly fall in the field of view of the laser scanner or the camera. We demonstrate successful application of our methodology by providing simulation studies and experimental results.

The inlier points extracted by our BnB approach can be used in any nonlinear optimization method to explicitly compute the extrinsic camera to laser transformation. The BnB method also provides a rough estimate of the extrinsic transformation which can be used for initializing the combined nonlinear optimization close to the optimal transformation, preventing it from local optimums.

II. PROBLEM FORMULATION

Consider a camera-laser setup attached rigidly together. Denote the camera-fixed coordinate frame and the laser-fixed coordinate frame by $\{c\}$ and $\{l\}$ respectively. For simplicity, we assume that only one checkerboard is present in the scene. Nevertheless, the method presented in this paper is directly applicable to the case where multiple checkerboards are present (see Remark 1). The representation of a point p in the camera frame and laser scanner frame are denoted by $p^c \in \mathbb{R}^3$ and $p^l \in \mathbb{R}^3$, respectively. One has

$$p^l = \Phi p^c + \Delta, \tag{1}$$

where $\Phi \in \mathrm{SO}(3)$ and $\Delta \in \mathbb{R}^3$, respectively, denote the rotation and translation from $\{c\}$ to $\{l\}$ and represent the extrinsic camera-laser calibration.
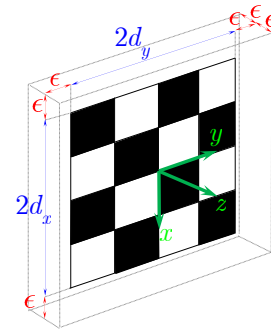

Fig. 1: The checkerboard coordinate frame is shown in green. The checkerboard dimensions are $2d_x$ and $2d_y$. A laser point is considered an inlier if it falls inside the dashed box around the checkerboard whose dimensions are $(2d_x + 2\epsilon) \times (2d_y + 2\epsilon) \times 2\epsilon$.

Assume that a calibration plane (e.g. a checkerboard) is placed in the environment such that it falls within the field of view of both the camera and the laser scanner. The laser scanner measures the 3D coordinate of the points of the environment in the laser-fixed frame. Some of these points fall in the plane of the checkerboard while others correspond to the points in the environment. For the purpose of this paper, we do not differentiate between 2D or 3D laser scanners as the theory presented here is applicable to both without any modification.

We assume the dimensions of the checkerboard are known and use standard image-based techniques to extract the location of intersection points in the image and find the camera intrinsic parameters as well as the pose parameters that relate the checkerboard's reference frame to the camera's. This task is easily accomplished using off-the-shelf software packages such as [23]. Changing the relative pose of the checkerboard with respect to (wrt.) the camera-laser setup (by moving either the checkerboard or the camera-laser setup), we collect n sets of laser scan points P_{ij} each contain m points ($i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$). Each set of these points corresponds to a scene in which a checkerboard with a known extrinsic transformation wrt. the camera frame is present. The main objective of this paper is to develop an algorithm that takes n sets of the laser scan points and their corresponding camera images and separates the laser points that correspond to the checkerboard (inliers) from those that belong to the background (outliers).

III. CHECKERBOARD EXTRACTION USING BNB

We propose a branch-and-bound method (BnB) [18], [20] for checkerboard extraction. Consider a coordinate frame attached to the center of the checkerboard with its z-axis normal to the checkerboard plane pointing outward (to the camera), its y-axis horizontal and its x-axis vertical, as depicted in Fig. 1. Denoting the dimensions of the checkerboard along its x and y direction by $2d_x$ and $2d_y$, a laser point is considered an inlier if it falls inside a box around the calibration board whose dimensions along the x, y, and z axes of the checkerboard are $2(d_x + \epsilon)$, $2(d_y + \epsilon)$, and 2ϵ, respectively, where $\epsilon > 0$ is a user defined threshold (see Fig. 1). For the checkerboard in the i-th image, denoted by N_i^x, N_i^y, and N_i^z vectors along the x, y, and z axes of the

checkerboard whose magnitude equals to the distance of the center of the camera frame to the $y - z$, $x - z$, and $x - y$ planes of the checkerboard coordinate frame, respectively. By detecting and processing the checkerboard corners, one can compute N_i^x, N_i^y, and N_i^z in the camera coordinate frame It is straight-forward to show that a laser point p_{ij}^l falls inside the box of Fig. 1 iff the following conditions hold [1].

$$|\bar{N}_i^{x\top} \Phi^\top (p_{ij}^l - \Delta) - \|N_i^x\|| < d_x + \epsilon, \tag{2a}$$

$$|\bar{N}_i^{y\top} \Phi^\top (p_{ij}^l - \Delta) - \|N_i^y\|| < d_y + \epsilon, \tag{2b}$$

$$|\bar{N}_i^{z\top} \Phi^\top (p_{ij}^l - \Delta) - \|N_i^z\|| < \epsilon, \tag{2c}$$

where $\bar{N} := \frac{N}{\|N\|}$ denotes the unit vector parallel to N. Inspired by [18], we propose the following objective function.

$$Q(\Phi, \Delta) = \sum_{i=1}^{n} \sum_{j=1}^{m} \lfloor |\bar{N}_i^{x\top} \Phi^\top (p_{ij}^l - \Delta) - \|N_i^x\|| < d_x + \epsilon \rfloor$$
$$\times \lfloor |\bar{N}_i^{y\top} \Phi^\top (p_{ij}^l - \Delta) - \|N_i^y\|| < d_y + \epsilon \rfloor$$
$$\times \lfloor |\bar{N}_i^{z\top} \Phi^\top (p_{ij}^l - \Delta) - \|N_i^z\|| < \epsilon \rfloor, \tag{3}$$

where $\lfloor . \rfloor$ is the indicator function. For a given transformation (Φ, Δ), any point p_{ij}^l that satisfies (2) is an inlier and contributes 1 to the summation in (3). Hence, Q represents the total number of the inliers for a given transformation. We wish to maximize Q over the transformation (Φ, Δ).

Algorithm 2 in the Appendix summarizes the BnB method used in this paper. We use the angle-axis representation to parametrize SO(3). A 3D rotation is represented as a vector whose direction and norm specify the axis and angle of rotation, respectively [17], [24]. This simplifies the search space of $\Phi \in$ SO(3) to a ball with a radius of π. We enclose this ball with a box with a half length of π, denoted by $\mathbb{B}(\pi)$. In practice, usually an initial estimate of the bound on the amplitude of the rotation Φ is available. This initial estimate can be used to initialize the search space of rotations in the queue with a smaller box $\mathbb{B}(\delta_R)$ with $\delta_R \leq \pi$. Similarly, the search space of translation is initialized with $\mathbb{B}(\delta_t)$ where δ_t represent an estimate of the maximum distance of the center of the camera from the center of the laser scanner.

The BnB method requires an upper bound for (3). Our main contribution is to propose an upper bound in Section IV (see Lemma 1). For branching, we simultaneously divide the boxes into eight sub-boxes, each with a half of the side length of the original box. Since we do the branching of the rotation and translation search spaces simultaneously at lines 7 and 8 of Algorithm 2, this branching yields the total of $8 \times 8 = 64$ new branches. Note that the inliers that satisfy the condition (2) are detected during evaluation of the objective function. Indexes of these inliers are returned as an output of the BnB method (see the line 21 of Algorithm 2).

IV. UPPER-BOUND FUNCTION

The key stage for employing the BnB method is to propose an upper-bound $\hat{Q}$ for (3). The following Lemma proposes an upper-bound and proves that it qualifies as a valid bounding function for BnB. Detailed derivations are given in the proof.

Lemma 1: Consider the boxes $\mathbb{B}_R$ and $\mathbb{B}_t$ with the centers $\Phi_c \in$ SO(3) and $\Delta_c \in \mathbb{R}^3$ and the half side length of δ_R and δ_t, respectively. Let

$$\hat{Q}(\mathbb{B}_R, \mathbb{B}_t) := \sum_{i=1}^{n} \sum_{j=1}^{m} \lfloor |\bar{N}_i^{x\top} \Phi_c^\top (p_{ij}^l - \Delta_c) - \|N_i^x\|| < d_x + \epsilon + \delta_{ij} \rfloor$$
$$\times \lfloor |\bar{N}_i^{y\top} \Phi_c^\top (p_{ij}^l - \Delta_c) - \|N_i^y\|| < d_y + \epsilon + \delta_{ij} \rfloor$$
$$\times \lfloor |\bar{N}_i^{z\top} \Phi_c^\top (p_{ij}^l - \Delta_c) - \|N_i^z\|| < \epsilon + \delta_{ij} \rfloor \tag{4}$$

where δ_{ij} is computed for each laser point p_{ij}^l as

$$\delta_{ij} = \|p_{ij}^l - \Delta_c\| \sqrt{2(1 - \cos(\sqrt{3}\delta_R))} + \sqrt{3}\delta_t. \tag{5}$$

The candidate upper-bound (4) satisfies

$$\hat{Q}(\mathbb{B}_R, \mathbb{B}_t) \geq \max_{\Phi \in \mathbb{B}_R, \ \Delta \in \mathbb{B}_t} Q(\Phi, \Delta). \tag{6}$$

Also, if $\mathbb{B}_R \times \mathbb{B}_t$ collapses to the single points (Φ, Δ), we have $\hat{Q}(\mathbb{B}_R, \mathbb{B}_t) = Q(\Phi, \Delta)$. $\square$

Proof of Lemma 1: Using the triangle inequality, we have

$$|\bar{N}^\top \Phi_c^\top (p_{ij}^l - \Delta_c) - \|N\|| \leq |\bar{N}^\top \Phi^\top (p_{ij}^l - \Delta) - \|N\||$$
$$+ |\bar{N}^\top \Phi_c^\top (p_{ij}^l - \Delta_c) - \bar{N}^\top \Phi^\top (p_{ij}^l - \Delta)|, \tag{7}$$

for any $N \in \mathbb{R}^3$. Again, applying triangle inequality yields

$$|\bar{N}^\top \Phi_c^\top (p_{ij}^l - \Delta_c) - \bar{N}^\top \Phi^\top (p_{ij}^l - \Delta)|$$
$$\leq |\bar{N}^\top \Phi^\top (\Delta - \Delta_c)| + |(\Phi_c \bar{N} - \Phi \bar{N})^\top (p_{ij}^l - \Delta_c)|$$
$$\leq \|\Phi \bar{N}\| \|\Delta - \Delta_c\| + \|\Phi_c \bar{N} - \Phi \bar{N}\| \|p_{ij}^l - \Delta_c\|. \tag{8}$$

Since $\Delta \in \mathbb{B}_t$ and Δ_c is the center of the box $\mathbb{B}_t$ with the half side length δ_t, we have $\|\Delta - \Delta_c\| \leq \sqrt{3}\delta_t$. Also, since $\Phi \in \mathbb{B}_R$ and Φ_c is the center of the box $\mathbb{B}_R$ with the half side length δ_R, and resorting to [25, equation (6)] and noting that $\|\bar{N}\| = 1$, we have $\|\Phi_c \bar{N} - \Phi \bar{N}\| \leq \sqrt{2(1 - \cos(\sqrt{3}\delta_R))}$. Combining these with (7) and (8) we have

$$|\bar{N}^\top \Phi_c^\top (p_{ij}^l - \Delta_c) - \|N\|| \leq |\bar{N}^\top \Phi^\top (p_{ij}^l - \Delta) - \|N\||$$
$$+ \sqrt{3}\delta_t + \|p_{ij}^l - \Delta\| \sqrt{2(1 - \cos(\sqrt{3}\delta_R))}. \tag{9}$$

Hence, by (5), $|\bar{N}^\top \Phi^\top (p_{ij}^l - \Delta) - \|N\|| \leq \bar{\epsilon}$ implies $|\bar{N}^\top \Phi_c^\top (p_{ij}^l - \Delta_c) - \|N\|| - \delta_{ij} \leq \bar{\epsilon}$ for any $\bar{\epsilon} \geq 0$. Consequently, $\lfloor |\bar{N}^\top \Phi^\top (p_{ij}^l - \Delta) - \|N\|| \leq \bar{\epsilon} \rfloor \leq \lfloor |\bar{N}^\top \Phi_c^\top (p_{ij}^l - \Delta_c) - \|N\|| \leq \bar{\epsilon} + \delta_{ij} \rfloor$. The above derivations are valid if one replaces N with N_i^x, N_i^y, or N_i^z and substitutes $\bar{\epsilon}$ for $d_x + \epsilon$, $d_y + \epsilon$, or ϵ, respectively. This proves (6). If $\mathbb{B}_R$ and $\mathbb{B}_t$ collapse to single points, we have $\Phi = \Phi_c$, $\Delta = \Delta_c$, and $\delta_R = \delta_t = 0$ (thus $\delta_{ij} = 0$). Substituting these into (4) yields $\hat{Q}(\mathbb{B}_R, \mathbb{B}_t) = Q(\Phi, \Delta)$ and completes the proof. ∎

A. Improving the tightness of the upper-bound

Tightness of the upper bound function (4) depends directly on the value of δ_{ij}. The less conservative value of δ_{ij} is used, the tighter upper bound $\hat{Q}$ is obtained and BnB finds the global optimal value with less iterations. In this section, we tighten the upper bound function proposed in Section IV.

Derivation of (5) relies on obtaining upper bounds of the inner products $|\bar{N}^\top \Phi^\top (\Delta - \Delta_c)|$ and $|(\Phi_c \bar{N} - \Phi \bar{N})^\top (p_{ij}^l -$
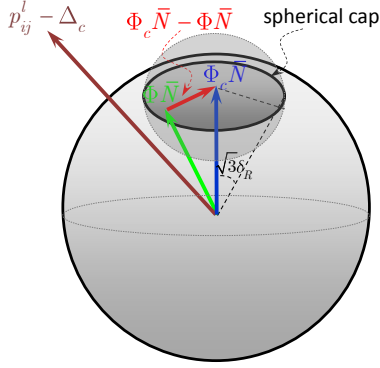
Fig. 2: The vector $\Phi\bar{N}$ belongs to the spherical cap whose axis of symmetry is $\Phi_c\bar{N}$. Computation of δ_{ij} via (5) involves obtaining an upper bound for the inner product $(\Phi_c\bar{N} - \Phi\bar{N})^\top(p_{ij}^l - \Delta_c)$ by enclosing the spherical cap with the small ball depicted in the figure. The tighter bound (11) obtains the upper bound directly on the spherical cap.

Algorithm 1 Fast computation of the extremums of $g(\Phi)$

Require: the normal of the calibration plane N, the laser point p_{ij}^l, the center of the box $\mathbb{B}_t$ given by Δ_c, the center and half side length of the box $\mathbb{B}_R$ given by Φ_c and δ_R.

1: Compute $\beta = \angle(\Phi_c\bar{N}, p_{ij}^l - \Delta_c)$.
2: **if** $\beta \leq \sqrt{3}\delta_R$ **then** $g_{\max} = \|p_{ij}^l - \Delta_c\|$.
3: **else**
4:    $g_{\max} = \|p_{ij}^l - \Delta_c\| \max(\cos(\beta - \sqrt{3}\delta_R), \cos(\beta + \sqrt{3}\delta_R))$.
5: **end if**
6: **if** $\beta \geq \pi - \sqrt{3}\delta_R$ **then** $g_{\min} = -\|p_{ij}^l - \Delta_c\|$.
7: **else**
8:    $g_{\min} = \|p_{ij}^l - \Delta_c\| \min(\cos(\beta - \sqrt{3}\delta_R), \cos(\beta + \sqrt{3}\delta_R))$.
9: **end if**
10: **return:** $g_{\max}$ and $g_{\min}$.

$\Delta_c)|$ over $(\Delta, \Phi) \in \mathbb{B}_t \times \mathbb{B}_R$. These upper bounds are obtained in (8) by bounding the inner products with the product of the magnitude of the involved vectors, without considering the angle between these vectors. This yields a reasonably tight result for the inner product $|\bar{N}^\top\Phi^\top(\Delta - \Delta_c)|$ as the vector $\Delta - \Delta_c$ belongs to the 3D box $\mathbb{B}(\delta_t)$ and inevitably there exist some Δ in the box for which the angle between $\Delta - \Delta_c$ and $\Phi_c N$ is zero. For the inner product $(\Phi_c\bar{N} - \Phi\bar{N})^\top(p_{ij}^l - \Delta_c)$, however, the angle between the involved vectors might not be zero (or even close to zero) since $\angle(\Phi_c\bar{N}, \Phi\bar{N}) \leq \sqrt{3}\delta_R$ and the vector $\Phi\bar{N}$ belongs to the spherical cap shown in Fig. 2 (see [17]). In fact, the upper bound (5) is obtained by bounding the spherical cap with a ball with the radius $\sqrt{2(1 - \cos(\sqrt{3}\delta_R))}$. Here, we obtain a tighter upper bound by maximizing the inner product directly on the spherical cap, taking into account the angle between $\Phi_c\bar{N} - \Phi\bar{N}$ and $p_{ij}^l - \Delta_c$. We have

$$|(\Phi_c\bar{N} - \Phi\bar{N})^\top(p_{ij}^l - \Delta_c)| = |\text{const} - (\Phi\bar{N})^\top(p_{ij}^l - \Delta_c)|. \quad (10)$$

where $\text{const} := (\Phi_c\bar{N})^\top(p_{ij}^l - \Delta_c)$. Hence, the particular value of $\Phi \in \mathbb{B}_R$ that maximizes (10) corresponds to the value that either minimizes or maximizes $g(\Phi) := (\Phi\bar{N})^\top(p_{ij}^l - \Delta_c)$. Algorithm 1 provides a fast approach for computing the maximum and minimum values of $g(\Phi)$, denoted by $g_{\max}$ and $g_{\min}$, respectively (proof of this algorithm is given in Appendix I). Using Algorithm 1, a tight upper bound for (10) is obtained as $|(\Phi_c\bar{N} - \Phi\bar{N})^\top(p_{ij}^l - \Delta_c)| \leq \max(|\text{const} - g_{\min}|, |\text{const} - g_{\max}|)$. In order to implement this new upper bound instead of the upper bound proposed in Section IV, one only needs to replace (5) with the following equation (the rest of the Algorithm 2 remains unchanged).

$$\delta_{ij} = \sqrt{3}\delta_t + \max(|\text{const} - g_{\min}|, |\text{const} - g_{\max}|). \quad (11)$$

Similar to (5), the new upper bound (11) still depends on the laser points p_{ij}^l and needs to be computed for each laser point separately. It is straight-forward to show that the tight upper bound $\hat{Q}$ based on (11) also satisfies the requirements of Lemma 1. We provide numerical comparison of the bound (5) and (11) in Section V (see Fig. 4).

Remark 1: Multiple checkerboards are sometimes used for camera-laser calibration [11]. In this case, detecting the inliers is harder as it is not known which one of the laser points in a given image-laser scan pair is associated with which one of the checkerboards. Assume that k_i checkerboards are visible in the i-th image and denote the normal vectors associated with the k-th checkerboard ($k \in \{1, \ldots, k_i\}$) by $N_{ik_i}^x$, $N_{ik_i}^y$, and $N_{ik_i}^z$. The number of observed checkerboards in each image can be different and it is allowed that some of the checkerboards do not fall within the field of view of the camera or laser scanner in some image-laser pairs. We extend the objective function (3) to

$$Q(\Phi, \Delta) =$$
$$\sum_{i=1}^{n}\sum_{j=1}^{m} \max_{k \in k_i}\Big(\lfloor|\bar{N}_{ik_i}^{x\top}\Phi^\top(p_{ij}^l - \Delta) - \|N_{ik_i}^x\|| < d_x + \epsilon\rfloor$$
$$\times \lfloor|\bar{N}_{ik_i}^{y\top}\Phi^\top(p_{ij}^l - \Delta) - \|N_{ik_i}^y\|| < d_y + \epsilon\rfloor$$
$$\times \lfloor|\bar{N}_{ik_i}^{z\top}\Phi^\top(p_{ij}^l - \Delta) - \|N_{ik_i}^z\|| < \epsilon\rfloor\Big). \quad (12)$$

Using the max operator where the correspondence between the measurements are unknown has been successfully practiced in the literature (see e.g. [18], [25]). For a given calibration (Φ, Δ), the objective function (12) choses the correspondence between the checkerboards and the laser scans that yields the maximum number of inliers. Note that extending the objective function to (12) does not change the structure of Algorithm 2 at all and the results of Lemma 1 still hold. One only needs to compute $Q(\Phi_c, \Delta_c)$ and $\hat{Q}(\mathbb{B}_R, \mathbb{B}_t)$ using (12) instead of (3). The optimum correspondence between the laser points and the checkerboards in each image is obtained at line 21 of Algorithm 2. $\square$

V. SIMULATION RESULTS

This section provides simulation studies demonstrating that the proposed BnB method is able to effectively extract the checkerboards even in challenging situations where some of the checkerboards are placed on the walls (hence do not create discontinuity in range measurements), or where some of the checkerboards either partially hit or are not hit at all by laser scans in some of the image-laser scan pairs.
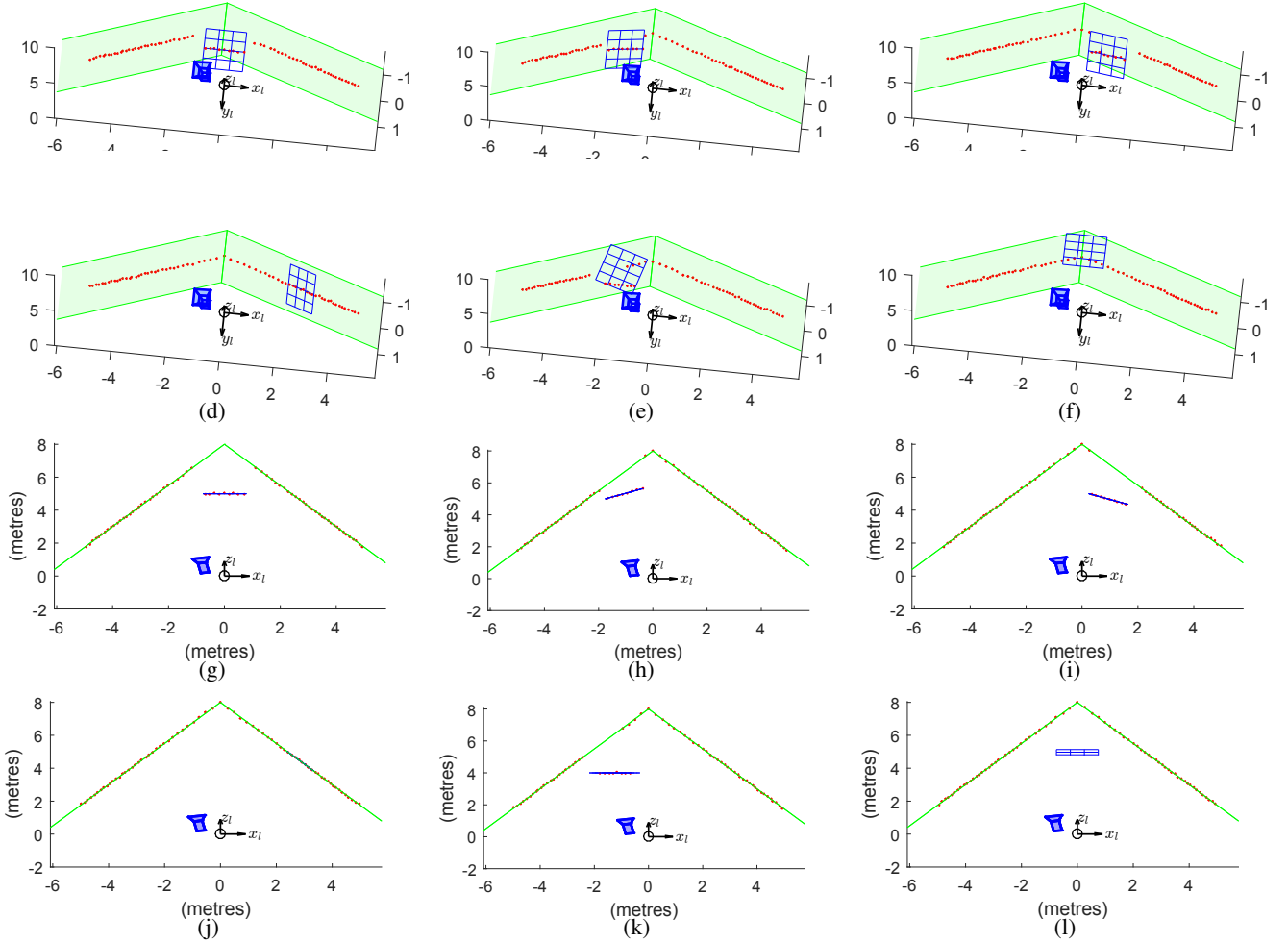
Fig. 3: Synthetic data used for the simulation example of Section V. The top six figures are 3D plots and the bottom six figures show their corresponding top views. The laser points are depicted by red dots, expressed in the laser coordinate frame depicted by a black coordinate frame whose axes are labeled as x_l, y_l, z_l. The camera is depicted by a blue camera sign, the checkerboard is depicted by a blue mesh surface, and the walls are shown in green. The checkerboard in Fig. 3d is placed on a wall preventing any discontinuity of the range measurements. The laser scans only hit a part of the checkerboard in Fig. 3e and they do not hit the checkerboard at all in Fig. 3f.
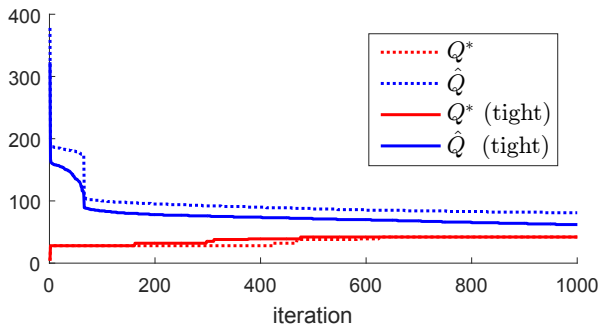

Fig. 4: Evolution of $Q^\star$ and the upper bound $\hat{Q}$ for the simulations of Section V. Dotted lines and solid lines, respectively, correspond to the case where the upper bound based on (5) and (11) are employed.

Fig. 5: Red dots are the inliers laser points identified by the BnB method (cf. Fig. 3g to Fig. 3l).

We choose the extrinsic rotation Φ to be a rotation of 10 degrees around the y-axis of the laser scanner and we choose the extrinsic translation $\Delta = [-0.75 \ -0.2 \ 0.5]^\top$. A $1.5(\text{m}) \times 1.5(\text{m})$ checkerboard is placed in 6 different positions in front of the camera and laser scanner, each time with a different orientat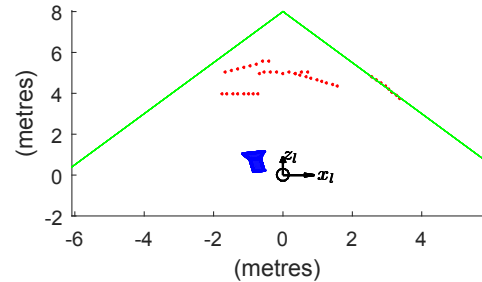ion (see Fig. 3). We consider a 2D laser scanner providing range measurements in a horizontal plane, every 2 degrees from -70 degrees to 70 degrees yielding a total of 71 measured laser points in each image. The laser-camera setup is placed in a triangular shaped room where the distance of the walls to the laser scanner is 5 meters and the intersection line of the walls is 8 meters away from the laser scanner. The laser points that do not hit the checkerboards will hit the walls. A uniformly distributed random noise between -2 centimeters and 2 centimeters is

added to the resulting range measurements.

The 6 synthetically generated sets are depicted in Fig. 3. To make the data more challenging for checkerboard extraction, in Fig. 3d, the checkerboard is placed exactly on the wall, preventing any discontinuity in the range measurements. Also, we choose the checkerboard orientation such that the laser scans only partially hit it in Fig. 3e and do not hit it at all in Fig. 3f. The total of $6 * 71 = 426$ laser points are measured, out of which 42 points correspond to the checkerboards (inliers) and the rest fall on the walls. Checkerboard normals represented in the camera frame are polluted by a uniformly distributed random rotation of maximum 1 degrees along each axis of the camera to model the camera noise. The resulting checkerboard normals along with the laser scan points are passed to Algorithm 2. The threshold ϵ is chosen to be 7 centimeters and the initial search space is chosen as $\mathbb{B}_R = \mathbb{B}(\frac{\pi}{12})$ and $\mathbb{B}_t = \mathbb{B}(1 \text{ meter})$.

Fig. 4 shows the evolution of $Q^\star$ and $\hat{Q}$ for the first 1000 iterations of BnB when they are examined at line 5 of Algorithm 2. Fig. 4 includes both the case where the original bound (5) or the tighter bound (11) is used in BnB. The upper bounds initially decay very fast, but their rate of decay decreases for large iterations (this is a typical characteristic of objective functions designed using the floor operator). As expected, $\hat{Q}$ computed using (11) decays faster than that of (5), showing that (11) indeed yields a tighter upper bound. The BnB method is able to find all of the inliers after 475 iterations when the tight upper bound is used and after 625 iterations when the original upper bound is used. The runtime is reported in Section VI. Observe that $\hat{Q}$ still decays at the 1000-th iteration and will eventually reach $Q^\star$ and the algorithm terminates. It is a common practice to terminate BnB algorithms after large enough number of inliers are detected and $Q^\star$ stops growing for a large enough consecutive iterations. Fig. 5 shows the inliers detected at line 21 of Algorithm 2 (for both the bound (5) and (11) demonstrating that all of the 42 inliers are successfully detected.

VI. Experimental results

In this Section, we verify the effectiveness of the proposed BnB method using real images and 2D laser scan dataset of [8][1]. The dataset contains 20 images together with their corresponding laser scans. Each laser scan contains 401 range measurements in the horizontal plane, measured every 0.25 degrees from -50 degrees to 50 degrees. A checkerboard is placed in a different place in each image. The images are processed using the camera calibration toolbox [23]. This toolbox extracts out the corners of the checkerboard, computes camera intrinsics and lens distortion parameters, and provides extrinsic transformation of each checkerboard plane wrt. to the camera frame. The output of [23] is used both in our BnB method and in the toolbox of [8]. Having the output of the camera calibration toolbox, it is

straight-forward to compute the checkerboard normals N_i^x, N_i^y, and N_i^z required by Algorithm 2. The dimensions of the checkerboard rectangles are provided in the dataset, but the dimensions of the calibration board itself (which are larger than the area of the checkerboard, see Fig. 7) are not provided. From the images, we estimate the dimensions of the calibration board to be at least $0.83(m) \times 0.83(m)$ and we use these values as $2d_x$ and $2d_y$ in Algorithm 2. The search space is initialized to $\mathbb{B}_R = \mathbb{B}(\frac{\pi}{18})$ and $\mathbb{B}_t = \mathbb{B}(0.5)$ and the inlier thresholds is chosen as $\epsilon = 0.1 \text{ meters}^2$.

Fig. 6 shows the extracted checkerboards after the first 1000 iterations of our BnB method with the tight bound (top plot) versus the checkerboards extracted by the toolbox of [8][3]. The total of 1038 inlier laser points are extracted by our method and 1036 inliers are detected by the toolbox of [8]. The extracted inlier laser points of both methods are almost identical. No ground truth inliers are given in the data set. Nevertheless, since the employed data has been verified to work efficiently with the toolbox of [8], the validity of the extracted inliers of the BnB method is verified. Fig. 7 shows a sample of the image-laser scan pairs. Some of the points near the edges of the checkerboard fall on the hand or clothes of the person holding the checkerboard, and are correctly classified as outliers by Algorithm 2. Also, since d_x and d_y are chosen smaller than the true dimensions of the checkerboard, occasionally one or two inlier laser points at the edge of the checkerboard are classified as outliers by our method. Same effect is observed in the results of [8]. This is not important in practice as those laser points are naturally more noisy and might not be very helpful in calibration [26].

On our machine equipped with a Core i7-4720HQ processor, the total 1000 iterations of the BnB takes 249 seconds while the toolbox of [8] extracts the checkerboards in less than 2 seconds. We emphasis, however, that the calibration is a one-off process and the run time of BnB is negligible compared to the data acquisition, justifying its generality, flexibility with practical conditions, and robustness gains. We emphasis that [8] is tailored to 2D laser scanners, necessarily requires stationary background, and relies on range discontinuities, all of which are satisfied in the employed dataset. Our BnB method does not impose any of these assumption and is applicable to 3D laser scanners as well.

VII. Conclusion

We formulate the checkerboard extraction as a combinatorial optimization problem with a clear cut objective function and we propose a branch-and-bound technique for optimizing the objective. The proposed BnB method is able to robustly extract the checkerboard in a diverse range of practical scenarios, including but not limited to where either

---

[1]Available online: http://www-personal.acfr.usyd.edu.au/akas9185/AutoCalib/

[2]The search space is initialized large enough to contain the true calibration parameters Φ and Δ. By estimating the calibration parameters via the toolbox of [8], we verified that the angle of rotation corresponding to Φ is less than 5 degrees and the amplitude of the transformation is less than 0.5 meters along each axis. Algorithm 2 works if a larger search space is chosen, although this increases the number of iterations.

[3]We adjusted the line line extraction thresholds of[8] to detect all of the checkerboards. Otherwise, one of the checkerboards would not be detected.

Fig. 8: The dashed red curve shows the intersection of the plane $\theta = 0$ (or $\theta = \pi$) with the spherical cap. The intersection of the plane with the boundary of the spherical cap corresponds to the points $(\theta = 0, \phi = \sqrt{3}\delta_R)$ and $(\theta = \pi, \phi = \sqrt{3}\delta_R)$.

2D or 3D laser scanners are used, multiple checkerboards are present in the scene, checkerboards are placed on the walls and no range discontinuity is associated with checkerboard edges, some of the checkerboards are only partially hit or are not hit at all by the laser scans, background changes from one scan to another, and multiple unwanted objects are present in the scene creating undesired range discontinuities. We demonstrate effective application of the proposed method via simulation and experimental studies.

APPENDIX I

Consider a coordinate frame whose center is at the origin of $\Phi_c\bar{N}$ and its z-axis is along $\Phi_c\bar{N}$ (see Fig. 8). Choose the x-axis of this coordinate frame such that it falls within the plane containing $\bar{N}$ and $p_{ij}^l - \Delta_c$. The y-axis is defined by $z \times x$. Any vector $\Phi\bar{N} \in \mathbb{B}_R$ is expressed in the xyz coordinate as $\Phi\bar{N} = [\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi]^\top$ where ϕ is the angle between the z-axis $\Phi\bar{N}$, and θ is the angle between the x-axis and the projection of $\Phi\bar{N}$ into the xy plane. Similarly, the expression of $p_{ij}^l - \Delta_c$ is given by $p_{ij}^l - \Delta_c = \|p_{ij}^l - \Delta_c\|[\sin\beta, 0, \cos\beta]^\top$ where $\beta := \angle(\Phi_c\bar{N}, p_{ij}^l - \Delta_c)$. We have

$$g(\Phi) = g(\theta, \phi) := (\Phi\bar{N})^\top(p_{ij}^l - \Delta_c)$$
$$= \|p_{ij}^l - \Delta_c\|(\sin\beta\sin\phi\cos\theta + \cos\beta\cos\phi). \quad (13)$$

Finding the value of Φ that maximizes or minimizes (13) boils down to maximizing/minimizing (13) wrt. θ and ϕ. Observe that $0 \le \theta \le 2\pi$ and $0 \le \phi \le \sqrt{3}\delta_R$. We have $\frac{\partial g(\theta, \phi)}{\partial\theta} = 0 \Rightarrow \sin\theta = 0 \Rightarrow \theta = 0$ or π. This implies that extremums of (13) occur on the intersection of the plane that includes $\Phi_c\bar{N}$ and $p_{ij}^l - \Delta_c$ (this plane is characterized by $\theta = 0$ or $\theta = \pi$) and the spherical cap of Fig. 2 (the intersection is depicted in Fig. 8). Replacing for $\theta = 0$ or $\theta = \pi$ in (13), we have $g(0, \phi) = \|p_{ij}^l - \Delta_c\|\cos(\beta - \phi)$ and $g(\pi, \phi) = \|p_{ij}^l - \Delta_c\|\cos(\beta + \phi)$. Hence, $\frac{\partial g(0, \phi)}{\partial\phi} = \|p_{ij}^l - \Delta_c\|\sin(\beta - \phi)$ and $\frac{\partial g(\pi, \phi)}{\partial\phi} = -\|p_{ij}^l - \Delta_c\|\sin(\beta + \phi)$. Observe that $\frac{\partial g(0, \phi)}{\partial\phi} = 0$ yields either $\phi = \beta$ or $\phi = \pi + \beta$. Since $0 \le \phi \le \pi$, only $\phi = \pi + \beta$ is valid. Similarly, $\frac{\partial g(\pi, \phi)}{\partial\phi} = 0$ yields either $\phi = \pi - \beta$ or $\phi = -\beta$ where
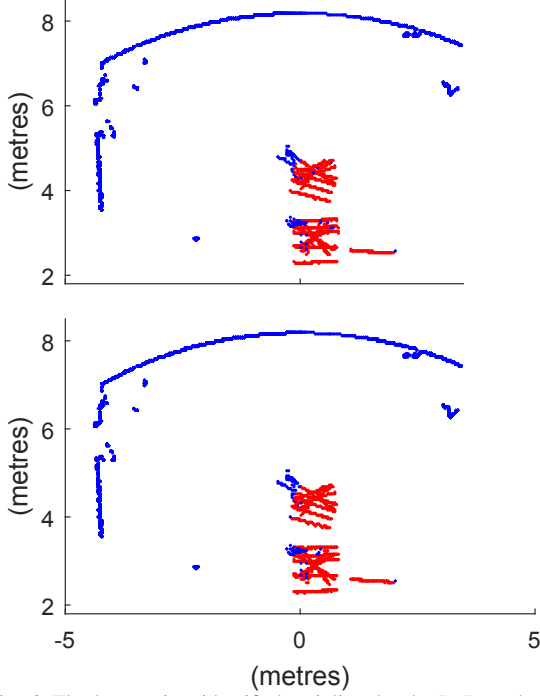

Fig. 6: The laser points identified as inliers by the BnB method (top plot) and by the toolbox of [8] (bottom plot) are colored red. The rest of the laser points (outliers) are colored blue. The laser scanner is placed at the origin.
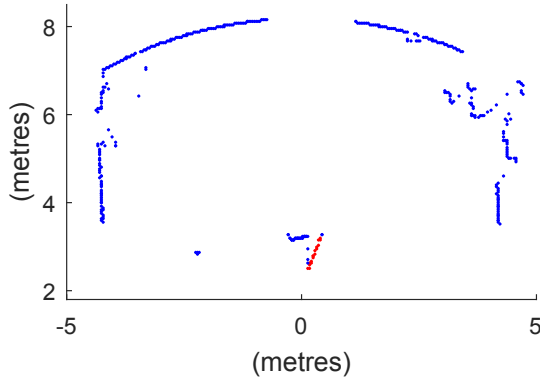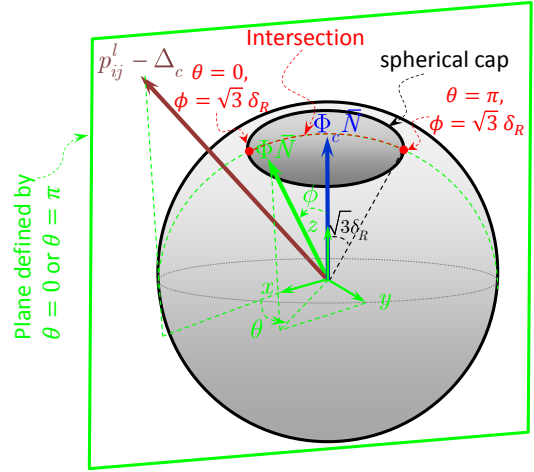

Fig. 7: A sample image (top) and its corresponding laser measurements (bottom). The inliers and outliers are identified by our BnB method are colored red and blue, respectively. Image and laser data from [8].

$\phi = -\beta$ is invalid. If either of $\phi = \beta$ or $\phi = \pi - \beta$ creates a rotation Φ that falls within the box $\mathbb{B}_R$, this rotation corresponds to an extremum of g. These correspond to the cases where either $\Phi\bar{N} = \frac{p_{ij}^l - \Delta_c}{\|p_{ij}^l - \Delta_c\|}$, yielding the maximum of (13) to be $g_{\max} = \|p_{ij}^l - \Delta_c\|$, or $\Phi\bar{N} = -\frac{p_{ij}^l - \Delta_c}{\|p_{ij}^l - \Delta_c\|}$, yielding the minimum of (13) to be $g_{\min} = -\|p_{ij}^l - \Delta_c\|$. Otherwise, since $g(\theta, \phi)$ is continuous on $\phi \in [0, \sqrt{3}\delta_R]$, the extremums might happen at the boundary point $\phi = \sqrt{3}\delta_R$, which yield the candidate extremums $g(0, \sqrt{3}\delta_R) = \|p_{ij}^l - \Delta_c\| \cos(\beta - \sqrt{3}\delta_R)$ and $g(\pi, \sqrt{3}\delta_R) = \|p_{ij}^l - \Delta_c\| \cos(\beta + \sqrt{3}\delta_R)$. Thus, in order to find the extremums, one should first check if either of $\phi = \beta$ or $\phi = \pi - \beta$ falls inside the box $\mathbb{B}_R$, yielding the maximum or minimum, respectively. Otherwise, the maximum/minimum is obtained from the boundary points. Algorithm 1 summarizes this method.

Algorithm 2 Branch-and-bound method for maximizing (3)

Require: ϵ, d_x, d_y, $\mathbb{B}(\delta_R)$, $\mathbb{B}(\delta_t)$, N_i^x, N_i^y, N_i^z, p_{ij}^l for $i = 1, \ldots, n, j = 1, \ldots, m$.

1: Initialize priority queue q, $\mathbb{B}_R \leftarrow \mathbb{B}(\delta_R)$, $\mathbb{B}_t \leftarrow \mathbb{B}(\delta_t)$.
2: Insert $(\mathbb{B}_R, \mathbb{B}_t)$ into q, $\Phi_c \leftarrow$ center of $\mathbb{B}_R$, $\Delta_c \leftarrow$ center of $\mathbb{B}_t$. $Q^\star \leftarrow Q(\Phi_c, \Delta_c)$.
3: **while** q is not empty **do**
4:     Find the element of q with the highest $\hat{Q}$ and remove this element from q and insert it into $(\mathbb{B}_R, \mathbb{B}_t)$.
5:     **if** $\hat{Q}(\mathbb{B}_R, \mathbb{B}_t) = Q^\star$ **then** terminate.
6:     **end if**
7:     $\mathbb{B}_{R_m}$, $m = 1, \ldots, 8 \leftarrow \text{branch}(\mathbb{B}_R)$.
8:     $\mathbb{B}_{t_n}$, $n = 1, \ldots, 8 \leftarrow \text{branch}(\mathbb{B}_t)$.
9:     **for** k=1 to 8 **do**
10:         **for** ℓ=1 to 8 **do**
11:             $\Phi_{ck} \leftarrow$ center of $\mathbb{B}_{R_k}$, $\Delta_{c\ell} \leftarrow$ center of $\mathbb{B}_{t_\ell}$.
                Compute $\hat{Q}(\mathbb{B}_{R_k}, \mathbb{B}_{t_\ell})$ and $Q(\Phi_{ck}, \Delta_{c\ell})$.
12:             **if** $\hat{Q}(\mathbb{B}_{R_k}, \mathbb{B}_{t_\ell}) > Q^\star$ **then**
13:                 Insert $(\mathbb{B}_{R_k}, \mathbb{B}_{t_\ell})$ into q.
14:                 **if** $Q(\Phi_{ck}, \Delta_{c\ell}) > Q^\star$ **then**
15:                     $Q^\star \leftarrow Q(\Phi_{ck}, \Delta_{c\ell})$,
                    $(\Phi^\star, \Delta^\star) \leftarrow (\Phi_{ck}, \Delta_{c\ell})$.
16:                 **end if**
17:             **end if**
18:         **end for**
19:     **end for**
20: **end while**
21: $inliers^\star \leftarrow$ index of the laser points that contribute 1 to the summation (3) in evaluating $Q(\Phi^*, \Delta^*)$.
22: **return** $\Phi^\star, \Delta^\star$, $inliers^\star$.

REFERENCES

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

[2] D. Scaramuzza, A. Harati, and R. Siegwart, "Extrinsic self calibration of a camera and a 3d laser range finder from natural scenes," in *Proc. IEEE/RSJ International Conf. Intelligent Robots and Systems*, 2007, pp. 4164–4169.

[3] F. Vasconcelos, J. P. Barreto, and U. Nunes, "A minimal solution for the extrinsic calibration of a camera and a laser-rangefinder," *IEEE Trans. pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2097–2107, 2012.

[4] A. Napier, P. Corke, and P. Newman, "Cross-calibration of push-broom 2d lidars and cameras in natural scenes," in *Proc. IEEE International Conf. Robotics and Automation*, 2013, pp. 3679–3684.

[5] J. Castorena, U. S. Kamilov, and P. T. Boufounos, "Autocalibration of lidar and optical cameras via edge alignment," in *IEEE International Conf. Acoustics, Speech and Signal Processing*, 2016, pp. 2862–2866.

[6] G. Pandey, J. R. McBride, S. Savarese, and R. M. Eustice, "Automatic targetless extrinsic calibration of a 3D lidar and camera by maximizing mutual information." in *Proc. AAAI Conf. on Artificial Intelligence*, 2012.

[7] R. Unnikrishnan and M. Hebert, *Fast extrinsic calibration of a laser rangefinder to a camera*. Carnegie Mellon University, 2005.

[8] A. Kassir and T. Peynot, "Reliable automatic camera-laser calibration," in *Proc. Australasian Conference on Robotics & Automation*, 2010.

[9] F. M. Mirzaei, D. G. Kottas, and S. I. Roumeliotis, "3D LIDAR–camera intrinsic and extrinsic calibration: Identifiability and analytical least-squares-based initialization," *The International Journal of Robotics Research*, vol. 31, no. 4, pp. 452–467, 2012.

[10] D. Herrera, J. Kannala, and J. Heikkilä, "Joint depth and color camera calibration with distortion correction," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 34, no. 10, pp. 2058–2064, 2012.

[11] A. Geiger, F. Moosmann, Ö. Car, and B. Schuster, "Automatic camera and range sensor calibration using a single shot," in *IEEE International Conf. Robotics and Automation*, 2012, pp. 3936–3943.

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

[13] G. Pandey, J. McBride, S. Savarese, and R. Eustice, "Extrinsic calibration of a 3D laser scanner and an omnidirectional camera," *IFAC Proceedings Volumes*, vol. 43, no. 16, pp. 336–341, 2010.

[14] D. Herrera, J. Kannala, and J. Heikkilä, "Accurate and practical calibration of a depth and color camera pair," in *International Conf. Computer analysis of images and patterns*, 2011, pp. 437–445.

[15] C. Mei and P. Rives, "Calibration between a central catadioptric camera and a laser range finder for robotic applications," in *Proc. IEEE Int. Conf. on Robotics and Automation*, 2006, pp. 532–537.

[16] 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.

[17] R. I. Hartley and F. Kahl, "Global optimization through rotation space search," *International Journal of Computer Vision*, vol. 82, no. 1, pp. 64–79, 2009.

[18] T. M. Breuel, "Implementation techniques for geometric branch-and-bound matching methods," *Computer Vision and Image Understanding*, vol. 90, no. 3, pp. 258–294, 2003.

[19] A. Parra Bustos, T.-J. Chin, A. Eriksson, H. Li, and D. Suter, "Fast rotation search with stereographic projections for 3D registration," *IEEE Trans. Pattern Analysis and Machine Intelligence*, (accepted in Dec. 2015).

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

[21] J. Heller, M. Havlena, and T. Pajdla, "A branch-and-bound algorithm for globally optimal hand-eye calibration," in *IEEE Conf. Computer Vision and Pattern Recognition*, 2012, pp. 1608–1615.

[22] Y. Seo, Y.-J. Choi, and S. W. Lee, "A branch-and-bound algorithm for globally optimal calibration of a camera-and-rotation-sensor system," in *IEEE International Conf. Computer Vision*, 2009, pp. 1173–1178.

[23] J.-Y. Bouguet, "Camera calibration toolbox for MATLAB," http://www.vision.caltech.edu/bouguetj/calib_doc/ (visited in Auguest 2016).

[24] A. Khosravian and M. Namvar, "Globally exponential estimation of satellite attitude using a single vector measurement and gyro," in *Proc. IEEE Conf. Decision and Control*, 2010.

[25] A. Parra Bustos, T.-J. Chin, and D. Suter, "Fast rotation search with stereographic projections for 3D registration," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2014, pp. 3930–3937.

[26] W. Boehler, M. B. Vicent, and A. Marbs, "Investigating laser scanner accuracy," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 34, pp. 696–701, 2003.