




Robust and Optimal Registration of Image Sets and Structured Scenes via Sum-of-Squares Polynomials

Danda Pani Paudel¹  · Adlane Habed² · Cédric Demonceaux³ · Pascal Vasseur⁴

Received: 25 April 2016 / Accepted: 18 August 2018 / Published online: 11 September 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

This paper addresses the problem of registering a known structured 3D scene, typically a 3D scan, and its metric Structure-from-Motion (SfM) counterpart. The proposed registration method relies on a prior plane segmentation of the 3D scan. Alignment is carried out by solving either the point-to-plane assignment problem, should the SfM reconstruction be sparse, or the plane-to-plane one in case of dense SfM. A Polynomial Sum-of-Squares optimization theory framework is employed for identifying point-to-plane and plane-to-plane mismatches, i.e. outliers, with certainty. An inlier set maximization approach within a Branch-and-Bound search scheme is adopted to iteratively build potential inlier sets and converge to the solution satisfied by the largest number of assignments. Plane visibility conditions and vague camera locations may be incorporated for better efficiency without sacrificing optimality. The registration problem is solved in two cases: (i) putative correspondences (with possibly overwhelmingly many outliers) are provided as input and (ii) no initial correspondences are available. Our approach yields outstanding results in terms of robustness and optimality.

Keywords 2D–3D registration · Structure-from-Motion · Polynomial Sum-of-Squares optimization

1 Introduction

The emergence of affordable 3D sensors and high quality 2D cameras has triggered a growing interest in combining both imaging modalities. 3D sensors provide faithful 3D scene models in the form of dense 3D point clouds while

images are used to extract texture information. High quality 3D models with mapped texture can be obtained so long as the 2D and 3D sensors are registered in a common reference frame. The two modalities are generally registered off-line and the 2D and 3D sensors kept rigidly attached during acquisition. Doing so, however, may prove impractical considering that suitable acquisition conditions for one sensor may not be adequate for the other (e.g. lighting conditions for cameras, surface orientation for 3D sensors, etc.). Some application-specific requirements (e.g. camera on a drone and a 3D scanner on a terrestrial vehicle) may altogether prohibit the sensors to be rigidly attached. When the 2D and 3D sensors are free to move with respect to each other, reliable methods for registering the two modalities, i.e. establishing inter-modality correspondences and estimating the rigid transformation aligning their reference frames, are highly desirable.

Structure-from-Motion (SfM) techniques compute 3D point coordinates from pixel correspondences across images. It is thus tempting to regard the problem of registering 3D and 2D sensors as that of aligning two 3D point sets: one set induced by the images and the other obtained from scanner measurements. Registering 3D point clouds is a well-studied problem. Most methods use the Iterative-Closest-Point (ICP)

Communicated by Josef Sivic.

✉ Danda Pani Paudel
paudel@vision.ee.ethz.ch

Adlane Habed
adlane.habed@icube.unistra.fr

Cédric Demonceaux
cedric.demonceaux@u-bourgogne.fr

Pascal Vasseur
pascal.vasseur@univ-rouen.fr

- ¹ Computer Vision Lab, ETH Zurich, Zurich, Switzerland
- ² ICube Laboratory, CNRS, University of Strasbourg, Strasbourg, France
- ³ Le2i Laboratory, CNRS, University of Bourgogne Franche-Comté, Dijon, France
- ⁴ Laboratoire d'Informatique, de Traitement de l'Information et des Systèmes, Normandie Univ, UNIROUEN, UNIHAVRE, INSA Rouen, LITIS, 76000 Rouen, France

algorithm (or its variants) (Bartoli and Castellani 2012; Fitzgibbon 2001; Rusinkiewicz and Levoy 2001) when the correspondences between two 3D point sets are not known. While ICP is a local method, Breuel (2003) investigated its global version using the Branch-and-Bound (BnB) algorithmic paradigm in several settings, mostly focusing upon the registration in 2D case. For 3D point sets registration, Enqvist et al. (2009) employed BnB to maximize the inlier set from the given point-to-point correspondences. Similar works also exist in the context of absolute pose estimation problem (Jurie 1999; Enqvist and Kahl 2008). A thorough insight about globally optimal solutions for geometric reconstruction problems can be found in Kahl and Henrion (2007). Recent work by Yang et al. (2013) (Go-ICP) provides a practical method for retrieving the globally optimal solution to same-scale point set registration. However, because SfM reconstructions suffer from a scale ambiguity, methods devised for registering same-scale data cannot be employed.

Most methods handling the scale ambiguity rely on establishing correspondences either between the 3D measurements obtained by both modalities or directly between scanned data and images (Liu and Stamos 2005; Christy et al. 1999; Ferraz et al. 2014). The sought transformation parameters are then obtained by either minimizing the registration loss function or maximizing the consensus set of inliers. Note that Random Sample Consensus (RANSAC) (Fischler and Bolles 1981) is the most widely used method for finding the set of maximum inliers. RANSAC, although not usually optimal, is a well established standard method for model estimation in the case of large amounts of outliers. Key to its success relies on its explicit inlier set maximization approach. Alternative approaches include robust estimation methods that minimize a robust loss function, such as M-estimator. Methods based on loss function minimization are more prone to outliers than their inlier-set-maximization counterparts (Bazin et al. 2013). Such methods may provide satisfactory results only when outliers are relatively few (Fraundorfer and Scaramuzza 2012).

Some approaches exploit scene knowledge or the Manhattan World assumption. For instance, methods have been devised based on line segment matching (Liu and Stamos 2005), target segmentation (Taneja et al. 2013), repeated patterns detection (Schindler et al. 2008), mutual information maximization (Mastin et al. 2009), and extended Chamfer matching (Zhang and Chen 2014). Registration methods relying on establishing correspondences may be undermined by unreliable visual feature descriptors. Alternative methods, not establishing initial correspondences, have also been proposed (Paudel et al. 2014; Du et al. 2007; Corsini et al. 2013; Plotz and Roth 2015). The methods in Paudel et al. (2014) and Du et al. (2007) use variants of the ICP algorithm and hence remain susceptible to partial scene overlap, scene occlusion, and high levels of outliers. The one in Corsini et al. (2013)

employs a RANSAC-based inlier set maximization in which the scale problem is handled by an extension of the 4-point congruent sets algorithm. A recent approach presented in Plotz and Roth (2015) computes the average gradient magnitude over all lighting directions under Lambertian shading. These gradients are then matched with the image gradient to obtain a coarse 2D–3D registration.

As far as maximizing the set of inliers is concerned, RANSAC is non-deterministic and provides no guarantee with respect to the optimality of its solution. Globally optimal inlier set maximization methods (Li 2009; Bazin et al. 2013) have recently been proposed for problems formulated through linear equations. Extensions to problems with nonlinear equations (Yang et al. 2014) is problem-specific, difficult and may result in much more complicated (possibly numerically intractable) mathematical formulations. Note that a variety of methods for solving systems of nonlinear polynomial equations exist. Some are based on Gröbner bases or homotopy continuation (Verschelde 1999; Habel et al. 2012). Others use Polynomial Sum-of-Squares (SoS) optimization (Schweighofer and Pinz 2008; Lasserre 2000; Parrilo 2000; Olsson et al. 2006; Chandraker et al. 2007; Chesi et al. 2002). However, such methods are dedicated to solving outlier-free systems and dealing with outliers is carried out through RANSAC.

In this paper, we address the problem of registering a 3D scan and a set of images of a structured scene captured by calibrated cameras. Our assumption is that the scene is structured in the sense that it can be segmented into, and represented by, planes (or planar patches). Such representation is compact (Borrmann et al. 2011) and can also be useful for scene knowledge-based refinement methods (Tamaazousti et al. 2011). The plane-based assumption is particularly valid when dealing with man-made environments, including (but not limited to) Manhattan World, urban and indoor scenes that are abundant with planes. In our approach, we seek the metric transformation relating the scene planes and the SfM-induced 3D structure. The SfM structure may either be represented by a sparse set of 3D points, obtained from sparse matching across images, or by planes should a sufficiently dense matching between images be obtained and the resulting point cloud segmented. When the SfM 3D structure is a sparse set of points, registration is carried out by establishing point-to-plane correspondences. Point-to-plane registration methods are known to perform better than their point-to-point counterparts (Segal et al. 2009). When the SfM-induced 3D structure consists of planes, registration is carried out by seeking plane-to-plane correspondences.

We rely on the fact that, under metric ambiguity, both point-to-plane and plane-to-plane assignments can be expressed as second degree polynomials in scaled-quaternion and translation parameters. Our approach aims at maximizing the set of inlier assignments with guaranteed optimality of

the consensus set. The consensus set maximization methods (Li 2009; Bazin et al. 2013) discussed above are not applicable because of the nonlinearity of the problem at hand. In our approach, we use the BnB algorithmic paradigm to explore the scaled-quaternion and translation parameter space. As in Li (2009) and Bazin et al. (2013), we rely on establishing optimistic and pessimistic sets of inlier assignments for pruning branches whose most optimistic sets are worse than the best pessimistic one. Our contribution is threefold:

- (i) We propose a novel modeling of the point-to-plane (possibly point-to-patch) and plane-to-plane correspondence problems. Our modeling is based on a rigorous Sum-of-Squares polynomial optimization theory and is used to derive new conditions to identify, with certainty, mismatched correspondences within parameter bounds. Our registration approach relies upon such conditions to build optimistic inlier assignment sets for given parameter bounds.
- (ii) We introduce SfM-specific constraints in our modeling, namely, a plane visibility criterion and optional vague constraints on the positions of the camera.
- (iii) Based on our modeling and constraints, we propose a globally optimal registration algorithm that maximizes the inlier set of either point-to-plane or plane-to-plane assignments in presence of putative correspondences along with its non-combinatorial counterpart in the absence of such correspondences.

Our paper is organized as follows: Sect. 2 provides the main results from polynomial SoS optimization theory that we exploit in our registration approach. The BnB paradigm is presented in Sect. 3. In Sect. 4, we present our working hypotheses along with the assignment polynomials arising in the point-to-plane and plane-to-plane alignment problems. In Sect. 5, we derive polynomial SoS conditions for identifying mismatches within given registration parameter bounds. In Sect. 6, we present our modeling of additional geometric constraints to handle the scale ambiguity, to exploit vague knowledge on camera locations, and to perform point-to-patch rather than point-to-plane registration. Plane visibility, vague camera locations, and (to a lesser extent) scaling are rather SfM-specific constraints. These, and the point-to-patch constraints, may be used to boost the search for mismatches. This leads us to the statement of our main result that we give in Sect. 7 along with the description of our BnB registration algorithm for point-to-plane (or point-to-patch) and plane-to-plane registration. The discussion about our method is presented in Sect. 8. The results of our experiments are summarized and discussed in Sect. 9. Section 10 concludes our work.

2 Polynomial Sum-of-Squares Theory

In this section, we present an overview of some important results in Polynomial Sum-of-Squares optimization theory. These results account for the main ingredients of our registration approach.

Definition 1 (SoS and PSD) Let $\mathbb{R}[x]$ be the ring of polynomials in n variables, $x = (x_1, x_2, \dots, x_n)$, with real-valued coefficients. A polynomial $f(x) \in \mathbb{R}[x]$ is

- Positive Semi-Definite (PSD) (or nonnegative) if $f(x) \geq 0$ for all $x \in \mathbb{R}^n$;
- Sum-of-Squares (SoS) if there exist polynomials $f_i(x) \in \mathbb{R}[x]$ such that $f(x) = \sum_i f_i(x)^2$.

A SoS is obviously always PSD and the converse is generally untrue. However, Hilbert (1888) proved that, for some classes of polynomials including quadratic ones, a polynomial is PSD if and only if it is SoS. Checking whether a polynomial is PSD is NP-hard (though decidable) while checking whether a polynomial is SoS is computationally tractable using Semi-definite Programming (SDP) and employing the so-called Gram matrix of the polynomial.

Definition 2 (Gram matrix Powers and Wörmann 1998) Consider a polynomial $f(x) \in \mathbb{R}[x]$ of degree $2d$. Let $\mathbf{Z}_d(x)$ be the vector of monomials of $f(x)$ up to monomials of degree d . The matrix \mathbf{G} such that $f(x) = \mathbf{Z}_d(x)^\top \mathbf{G} \mathbf{Z}_d(x)$ is a Gram matrix of $f(x)$.

Theorem 1 (Choi et al. 1995; Powers and Wörmann 1998) A polynomial $f(x) \in \mathbb{R}[x]$ of degree $2d$ is SoS if and only if there exists a real symmetric positive semi-definite Gram matrix of $f(x)$.

Note that since odd-degree polynomials cannot be SoS, only even-degree polynomials are concerned by such test. Checking for the existence of a positive semi-definite Gram matrix \mathbf{G} boils down to solving a Linear Matrix Inequality (LMI) feasibility problem. LMI feasibility can be efficiently solved using the interior-point algorithm (Boyd and Vandenberghe 2004). Theorem 1 allows us to check whether a polynomial $f(x)$ is nonnegative for every $x \in \mathbb{R}^n$. One is often interested in checking whether $f(x)$ is nonnegative in a semi-algebraic set \mathcal{K} defined by polynomials $g_i(x) \in \mathbb{R}[x]$ such that

$$\mathcal{K} = \{x \in \mathbb{R}^n : g_i(x) \geq 0, i = 1 \dots m\}. \quad (1)$$

This can be answered via the so-called *Positivstellensatz* (Psatz) (Parrilo 2000). The Psatz states that $f(x)$ is nonnegative on \mathcal{K} if and only if there exist SoS polynomials $\sigma_v(x)$ such that

$$f(x) = \sum_{v \in \{0,1\}^m} \sigma_v(x) g_1(x)^{v_1} g_2(x)^{v_2} \dots g_m(x)^{v_m}. \quad (2)$$

Exploiting Psatz is difficult and may turn numerically intractable in practice because (2) requires 2^m SoS σ_i polynomials. Putinar (1993) provides a much simpler Psatz under Archimedean conditions on the so-called quadratic module.

A module is a well-studied algebraic structure, like a vector space, defined over a ring rather than a field. Let the set of SoS polynomials be $\mathcal{S}[x] = \{g_i(x)\}_{i=1}^m$, defining the semi-algebraic set \mathcal{K} . Then, we are interested in the modules defined over the semi-ring of $\mathcal{S}[x]$. A semi-ring is an algebraic structure similar to a ring, but there is no requirement of additive inverse of each element. Then, a subset of ring $\mathbb{R}[x]$ is a module over $\mathcal{S}[x]$, if it is closed under addition and scalar multiplication. In fact, the *quadratic module* is a cone of sum-of-squares—which is also a module over $\mathcal{S}[x]$ —defined as follows:

Definition 3 (*Quadratic module* Wagner 2009) The quadratic module $M(g) = M(g_1, \dots, g_m) \subset \mathbb{R}[x]$ of polynomials $g_1(x), g_2(x), \dots, g_m(x)$ is the set

$$M(g) = \left\{ \sigma_0(x) + \sum_{i=1}^m \sigma_i(x)g_i(x) : \text{each } \sigma_i \text{ is SoS} \right\}. \quad (3)$$

Definition 4 (*Archimedean* Wagner 2009) The quadratic module $M(g)$ of polynomials $g_1(x), g_2(x), \dots, g_m(x)$ is Archimedean if $N - \|x\|^2 \in M(g)$ for some $N \in \mathbb{N}$.

The term Archimedean can also be related to the Archimedean property of ordered or normed groups, and fields. Intuitively, it is the property of having no infinitely large elements therefore ensuring an upper bound. This property appears as Axiom V of Archimedes’ “On the Sphere and Cylinder”. In our case, the norm of the variables defining module $M(g)$ is bounded by the upper bound \sqrt{N} for some $N \in \mathbb{N}$.

Theorem 2 (Putinar’s Positivstellensatz Putinar 1993) Assume the quadratic module $M(g)$ is Archimedean. If $f(x) > 0$ on \mathcal{K} (defined by (1)), then $f(x) \in M(g)$.

In the special case of degree 2 polynomials, we further relax the Putinar’s Positivstellensatz during positivity test. This is carried out by using only positive scalar coefficients instead of SoS polynomials, with the help of the following lemma.

Lemma 1 (Finsler’s 1936/37) Let y be a vector, Q a symmetric matrix, B a rectangular matrix—all real-valued and of appropriate dimensions—and $\gamma > 0$ a scalar. The following statements are equivalent:

- (i) $y^T Q y > 0, \forall y \neq 0 : B y = 0$.
- (ii) There exists $\gamma : Q + \gamma B^T B > 0$.

Finsler’s lemma allows to convert the problem of checking the sign of a quadratic form over a subspace into solving a LMI problem.

3 Branch-and-Bound Search

We perform the BnB search to obtain the parameter $x \in \mathbb{R}^n$ that maximizes the consensus among a given set of polynomials parametrized by x . While doing so, our BnB algorithm performs a hierarchical discretization of the optimization variables via a dynamically-built search-tree. The process of hierarchical discretization is also known as branching. The branching process explores the parameter space with the help of the optimistic and pessimistic number of inliers for each branch, seeking an ϵ -gap parameter interval containing the globally optimal solution.

3.1 Dynamic Tree Construction

Every node of the dynamic tree represents a closed convex subspace of the optimization variables. It is defined by the lower and upper bounds of the variables in the form of two vectors \underline{x} and \bar{x} in \mathbb{R}^n , respectively. The lower and upper bound entries \underline{x}_i and \bar{x}_i are such that $\underline{x}_i \leq \bar{x}_i$ for all $i = 1, \dots, n$. Thus a node is defined by the variable intervals $[\underline{x}, \bar{x}]$. $x \in [\underline{x}, \bar{x}]$ means that each entry x_i of x is such that $\underline{x}_i \leq x_i \leq \bar{x}_i$.

Starting from a known interval, say $\mathcal{B}_0 = [\underline{x}_{init}, \bar{x}_{init}]$, the tree is constructed by recursively dividing the interval into two or more tighter intervals. If \mathcal{B}_0 is divided into $\mathcal{B}_k, k = 1, \dots, p$ intervals, it must satisfy $\mathcal{B}_k \subset \mathcal{B}_0$ for all $k = 1, \dots, p$ and $\mathcal{B}_0 = \mathcal{B}_1 \cup \mathcal{B}_2 \dots \cup \mathcal{B}_p$.

We illustrate the tree construction process with an example of two variables, say $x = (x_1, x_2)^T$, scenario. Let a_0 be the first node representing the interval \mathcal{B}_0 . The interval \mathcal{B}_0 is subdivided into smaller intervals by breaking each variable bounds into two. Two variables, each with two bounds produce four combinations of smaller intervals, which are represented by first level nodes a_1, a_2, a_3 , and a_4 . The complete tree construction process is then carried out through similar branching in a recursive manner. Fig. 1(left) provides the graphical illustration of two levels of branching. The nodes created in this fashion are dynamically stored in a tree structure as shown in Fig. 1(right). In case of higher dimensional search space, trees are constructed in a very similar fashion. However, the number of branching per node can vary as per convenience.

3.2 Search Method

The Branch-and-Bound algorithmic paradigm relies on finding the optimistic and pessimistic number of inliers within

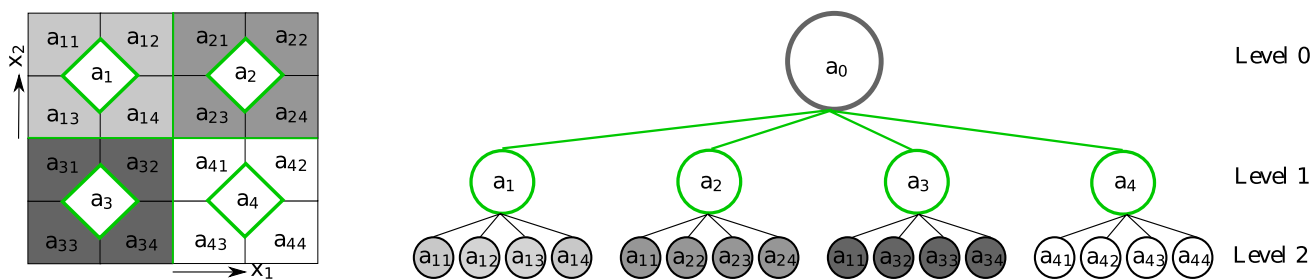


Fig. 1 Left: hierarchical branching in two dimensional space. The subspaces represented by nodes a_i and a_{ij} are the results of, respectively, the first and second levels of branching. Right: results of two levels of branching represented as by a tree structure (Color figure online)

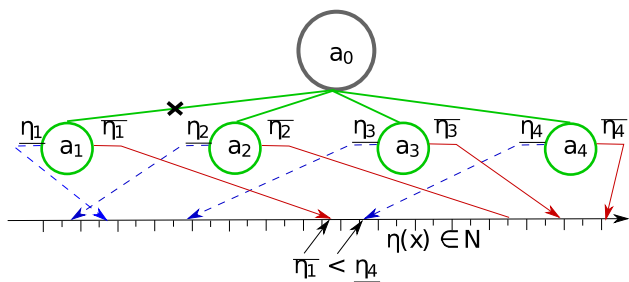


Fig. 2 Branch-and-Bound pruning based on bounds on the number of inliers. Node a_1 is pruned because its optimistic number of inliers is smaller than the pessimistic number of inliers of node a_4 (Color figure online)

each interval of variables. Any local solution, obtained using as a starting point a randomly picked sample from the given interval, can serve to determine a pessimistic number of inliers. Therefore, the most challenging part is to efficiently determine a tight optimistic number of inliers. An efficient method to estimate such optimistic number of inliers for our registration problem is given in Sect. 7. Here, we discuss the BnB process with an example while assuming that the optimistic and pessimistic inliers can be estimated for any given interval.

Consider the inliers measure $\eta(x) \in \mathbb{N}$ and $\underline{\eta}, \bar{\eta}$, respectively, the pessimistic and optimistic numbers of inliers estimated in $[x, \bar{x}]$. We represent the node-specific bounds using the subscript corresponding to that node. For the first level nodes, the mapping of both bounds to the axis with the number of inliers is shown in Fig. 2. One can observe from this diagram that the pessimistic number of inliers of node a_4 is greater than the optimistic number of inliers of a_1 (i.e. $\bar{\eta}_1 < \underline{\eta}_4$). Therefore, the solution corresponding to the maximum number of inliers cannot lie in the subspace represented by node a_1 . As a result, a_1 can be safely pruned. In fact, if the pessimistic number of inliers of any node is greater than the optimistic number of inliers of any other node, then nodes with smaller optimistic number of inliers can always be pruned. This process can be applied repeatedly in a recursive manner until the globally optimal number of inliers is reached. The algorithm terminates when the max-

imum number pessimistic inliers (obtained so far) becomes equal to the maximum optimistic inliers among all remaining nodes. The complete BnB process seeking the optimal number of inliers η^* is summarized in Algorithm 1. In our case, we also return the parameter x^* corresponding to η^* .

Algorithm 1 Branch-and-Bound Search

```

Input:  $\mathcal{B}_0$ 
Global variable:  $\eta$ 
1:  $\underline{\eta} = \text{computePessimisticInliers}(\mathcal{B}_0)$  ▷ Initialization
2:  $\eta^* = \text{processNodeBnB}(\mathcal{B}_0, \underline{\eta})$  ▷ Recursive function

3: function PROCESSNODEBNB( $\mathcal{B}, \eta$ )
4:    $\underline{t} = \text{computePessimisticInliers}(\mathcal{B})$ 
5:    $\bar{t} = \text{computeOptimisticInliers}(\mathcal{B})$ 
6:   if  $\underline{t} > \eta$  then
7:      $\eta \leftarrow \underline{t}$  ▷ Inliers update
8:   end if
9:   if ( $\bar{t} \leq \eta$ ) then ▷ BnB stopping criteria
10:    return  $\eta$ 
11:  else
12:    ( $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_k$ ) = divideBranches( $\mathcal{B}$ )
13:    for  $i = 1, 2, \dots, k$  do
14:      return processNodeBnB( $\mathcal{B}_i, \eta$ )
15:    end for
16:  end if
17: end function
    
```

4 Assignment Polynomials

In this paper, we consider a set of two or more calibrated cameras observing a scene consisting of a set \mathcal{P} of at least four distinct planes in general positions. The scene has been scanned by a 3D sensor and segmented into these planes. A plane $\Pi \in \mathcal{P}$ is given by its normal 3-vector π and signed distance to the origin d . We also consider the set \mathcal{Y} of seven or more points (lying on at least four distinct scene planes) whose projections are matched across two or more cameras. We distinguish two working hypotheses depending on whether image correspondences are sparse or dense:

- when the set \mathcal{Y} is sparse, the SfM-induced (Hartley and Zisserman 2004) 3D points are likewise sparse. Each reconstructed point is then represented by a $y \in \mathbb{R}^3$ of Cartesian coordinates;
- should \mathcal{Y} be sufficiently dense, the SfM-induced point cloud may further be segmented into a set of planes \mathcal{P}_r . Each such plane is represented by its normal 3-vector π_r and distance to the origin d_r .

The coordinates of the SfM-reconstructed points and/or planes and those of the scanned scene planes are represented in two distinct reference frames. The two representations of the scene also differ by a generally unknown scale factor. Consequently, the transformation aligning the SfM-reconstructed scene (whether represented by points or by planes) and the scanned scene is represented by a 3×3 scaled-rotation matrix Q and a translation 3-vector t . A quaternion representation with no enforcement of unit quaternion $q = (z \ u \ v \ w)^T$ is used to represent the scaled-rotation matrix Q :

$$Q = \begin{bmatrix} z^2 + u^2 - v^2 - w^2 & 2uv - 2wz & 2uw + 2vz \\ 2uv + 2wz & z^2 - u^2 + v^2 - w^2 & 2vw - 2uz \\ 2uw - 2vz & 2vw + 2uz & z^2 - u^2 - v^2 + w^2 \end{bmatrix}.$$

Aligning the SfM and scanned representations consists in finding Q and t that together map the 3D SfM-reconstructed points, or alternatively the SfM-induced planes, to their corresponding planes in the scanned scene. The problem of finding the correct matches between the 3D SfM-reconstructed points and scanned planes is referred to as the “point-to-plane” assignment problem. That of finding the correct match between the SfM and scanned planes is referred to as the plane-to-plane assignment problem. In both cases, the assignments are described by polynomials in the unknown entries of Q and t and which we denote by a vector $x \in \mathbb{R}^7$ such that $x = (q^T, t^T)^T$.

Point-to-Plane Assignment Polynomials When relying on point-to-plane assignments for registering the image and scanner modalities, we consider $\mathcal{A} \subset \mathcal{Y} \times \mathcal{P}$ as a set of putative point-to-plane assignments (\times refers to the cartesian product) and $a = (Y, \Pi) \in \mathcal{A}$ is one such assignment. The polynomial $f^a(x)$ in $\mathbb{R}[x]$ induced by a is given by:

$$f^a(x) := \pi^T(Qy + t) - d. \tag{4}$$

If x is the true registration parameter vector, then for every correct assignment $a \in \mathcal{A}$, $f^a(x) = 0$. Note that the residual of polynomial $f^a(x)$ is the orthogonal distance from point Y to plane Π for the parameter x . Accounting for scanning, reconstruction and round-off errors, this residual must lie within a reasonable margin for an assignment to be considered an inlier. Given $\xi \geq 0$, a sufficiently small predefined distance threshold, inlier assignments ought to satisfy

$$|f^a(x)| \leq \xi \text{ where } \xi \geq 0. \tag{5}$$

Note that ξ representing a geometric error, it is fairly easy to set. We are interested in finding x that maximizes the number of assignments by assessing the exact point-to-plane distance measure.

Plane-to-Plane Assignment Polynomials When carrying out plane-to-plane registration, the set of putative assignments \mathcal{A} represents a subset of $\mathcal{P}_r \times \mathcal{P}$. If the euclidean point Y relates to the SfM reconstructed point Y_r by $y_r = Qy + t$, then the Euclidean plane Π and reconstructed plane Π_r are related by,

$$\begin{bmatrix} \pi_r \\ -d_r \end{bmatrix} = \delta \begin{bmatrix} Q^T & 0 \\ t^T & 1 \end{bmatrix} \begin{bmatrix} \pi \\ -d \end{bmatrix}, \tag{6}$$

such that $\pi_r^T y_r - d_r = \pi_r^T(Qy + t) - d_r = 0$ is satisfied. Now, after introducing the quaternion terms, the assignment $a = (\Pi_r, \Pi) \in \mathcal{A}$ can be described by a 4-vector of degree 2 polynomials

$$f^a(x) := q^T q \begin{bmatrix} \pi_r \\ -d_r \end{bmatrix} + \delta \begin{bmatrix} Q^T & 0 \\ t^T & 1 \end{bmatrix} \begin{bmatrix} \pi \\ -d \end{bmatrix} \tag{7}$$

Should x be the true registration parameter vector, all four polynomials simultaneously vanish for some value of $\delta = \pm 1$. Similar to (5), accounting for errors, an inlier assignment must satisfy

$$|f_i^a(x)| \leq \xi \text{ for } i = 1, 2, 3, 4 \tag{8}$$

where $f^a(x) = (f_1^a(x), f_2^a(x), \dots, f_4^a(x))$. However, this time, the residuals do not correspond to a geometric distance but are rather of an algebraic nature. The threshold ξ in this case is hard to set. Nevertheless, a useful threshold can generally be obtained empirically such as trial-and-error. The work herein assumes such threshold is given.

5 Polynomial SoS Assignment Conditions

Our goal is to simultaneously estimate the registration parameters x and associated set of correct assignments. The problem can be solved by considering either point-to-plane or plane-to-plane assignments. Note, however, that we voluntarily do not distinguish between the cases in which the initial set \mathcal{A} is a point-to-plane putative assignments set from that in which it is a plane-to-plane one. Unless stated otherwise, $f^a(x)$ is always considered as a vector of polynomials of appropriate dimension p : a 1-vector with $p = 1$ to represent the single polynomial (4) induced by a point-to-plane putative assignment or, with $p = 4$, a 4-vector of polynomials (7) induced by a plane-to-plane assignment. We use $f_i^a(x)$ to refer to the i^{th} entry of $f^a(x)$, $1 \leq i \leq p$.

We may assume, for the sake of clarity of the exposition, that, in the case of plane-to-plane assignments, the value of $\delta = \pm 1$ in (4) is known. It will be made clear, further in this section (see Result 3), how the two cases are handled in our method.

Our registration approach is based on the BnB algorithmic paradigm and branching is carried out on the space of registration parameters \mathbf{x} , as discussed in Sect. 3. At each iteration, we are given parameter interval $[\underline{\mathbf{x}}, \bar{\mathbf{x}}]$. We probe this interval for potentially correct assignments by attempting to solve, for each assignment, the following problem:

Problem 1 For given $a \in \mathcal{A}$ and threshold $\xi \geq 0$, is there a vector $\mathbf{x} \in [\underline{\mathbf{x}}, \bar{\mathbf{x}}]$ such that $|f_i^a(\mathbf{x})| \leq \xi$ for all $i = 1 \dots p$?

In other words, through Problem 1, one would like to know whether all polynomials in $f^a(\mathbf{x})$ are within the desired margin for some \mathbf{x} within the considered bounds. The assignment would then qualify as a potential inlier, i.e. a possibly correct assignment, within these bounds. This is however difficult to answer. One alternative is to consider the following problem:

Problem 2 For given $a \in \mathcal{A}$ and threshold $\xi \geq 0$, is there a polynomial $f_i^a(\mathbf{x})$ in vector $f^a(\mathbf{x})$ satisfying $|f_i^a(\mathbf{x})| > \xi$ for every $\mathbf{x} \in [\underline{\mathbf{x}}, \bar{\mathbf{x}}]$?

If the question of Problem 2 is answered in the affirmative, the one of Problem 1 is answered in the negative: i.e. there exists no $\mathbf{x} \in [\underline{\mathbf{x}}, \bar{\mathbf{x}}]$ with which $|f_i^a(\mathbf{x})| \leq \xi$ for all $i = 1, \dots, p$. In such case, the assignment a is definitely an outlier, i.e. incorrect assignment, within the bounds. Otherwise, it is a potential inlier.

One can rely on Putinar’s Theorem 2 to solve Problem 2. To do so, assume we are given a set of polynomials $g_k(\mathbf{x})$ whose quadratic module $M(g)$ is Archimedean: if, for a scalar λ_i^a such that $-1 \leq \lambda_i^a \leq +1$, the polynomial $\lambda_i^a f_i^a(\mathbf{x}) > \xi$ for all $\mathbf{x} \in \mathcal{K} = \{\mathbf{x} \in \mathbb{R}^7 : g_k(\mathbf{x}) \geq 0, k = 1 \dots m\}$, then $\lambda_i^a f_i^a(\mathbf{x}) \in M(g)$. Hence, SoS polynomials $\sigma_k(\mathbf{x})$ and scalar $-1 \leq \lambda_i^a \leq +1$ exist such that:

$$\lambda_i^a f_i^a(\mathbf{x}) - \xi - \sum_{k=1}^m \sigma_k(\mathbf{x})g_k(\mathbf{x}) \text{ is SoS.} \tag{9}$$

Note that, in general, if (9) is satisfied, then $\lambda_i^a f_i^a(\mathbf{x})$ may not necessarily be positive in \mathcal{K} since \mathcal{K} could possibly be empty. However, so long as \mathcal{K} is not empty and $\sigma_k(\mathbf{x})$ SoS polynomials and scalar λ_i^a can be found, one is guaranteed that $\lambda_i^a f_i^a(\mathbf{x}) > \xi$ everywhere in \mathcal{K} since $\sum_{k=1}^m \sigma_k(\mathbf{x})g_k(\mathbf{x}) > 0$ in \mathcal{K} . Moreover, if $-1 \leq \lambda_i^a \leq +1$, then this guarantees that either $f_i^a(\mathbf{x}) > \xi$ or $-f_i^a(\mathbf{x}) > \xi$ everywhere in \mathcal{K} : assignment a would be an outlier.

There are two main pending issues before one is able to use (9). First, one needs to find a set of polynomials $g_k(\mathbf{x})$,

representative of the parameter intervals, whose quadratic module $M(g)$ is Archimedean. Second, it is so far unclear how the $\sigma_k(\mathbf{x})$ SoS polynomials can be found. Let us explore now the first of these issues. Note that the Archimedean property is a matter of representation and the quadratic module of the set constructed from the linear interval constraints $x_k - \underline{x}_k \geq 0$ and $\bar{x}_k - x_k \geq 0$ is not Archimedean. In the following, we show that quadratic polynomial inequalities derived from such bound constraints yield an Archimedean quadratic module.

Proposition 1 Consider the polynomials $g_k(\mathbf{x}) = (x_k - \underline{x}_k)(\bar{x}_k - x_k)$, $k = 1 \dots 7$. The quadratic module $M(g)$ of these polynomials is Archimedean.

Proof As per Definition 4, for $M(g)$ to qualify as Archimedean, one ought to demonstrate that $N - \sum_{k=1}^n x_k^2 \in M(g)$ for some $N \in \mathbb{N}$. Definition 3 indicates that, for $N - \sum_{k=1}^n x_k^2$ to be in $M(g)$ with the $g_k(\mathbf{x})$ polynomials of this proposition, there must exist SoS polynomials $\sigma_0(\mathbf{x})$ and $\sigma_k(\mathbf{x})$, $k = 1 \dots 7$, such that

$$N - \sum_{k=1}^n x_k^2 = \sigma_0(\mathbf{x}) + \sum_{k=1}^7 \sigma_k(\mathbf{x})(x_k - \underline{x}_k)(\bar{x}_k - x_k). \tag{10}$$

In particular, because $\sigma_0(\mathbf{x})$ must be SoS, the proof boils down to establishing the existence of SoS $\sigma_k(\mathbf{x})$, $k = 1 \dots 7$ that can turn $\sigma_0(\mathbf{x})$, of the form

$$\sigma_0(\mathbf{x}) = N - \sum_{k=1}^7 x_k^2 - \sum_{k=1}^7 \sigma_k(\mathbf{x})(x_k - \underline{x}_k)(\bar{x}_k - x_k), \tag{11}$$

into a SoS polynomial. The polynomial $\sigma_0(\mathbf{x})$ can be rewritten as

$$\sigma_0(\mathbf{x}) = N - \sum_{k=1}^7 \left(x_k^2 + \sigma_k(\mathbf{x})(x_k - \underline{x}_k)(\bar{x}_k - x_k) \right). \tag{12}$$

Let us first show that a PSD, not necessarily SoS, $\sigma_0(\mathbf{x})$ exists. To do so, we thus seek $\sigma_k(\mathbf{x})$, $k = 1 \dots 7$, for which there exists N satisfying

$$N \geq \max_{\mathbf{x}} \left(\sum_{k=1}^7 \left(x_k^2 + \sigma_k(\mathbf{x})(x_k - \underline{x}_k)(\bar{x}_k - x_k) \right) \right). \tag{13}$$

Such N exists if the polynomial argument of $\max(\cdot)$ is concave. Observe that using zero-degree SoS polynomials σ_k that are independent from \mathbf{x} , i.e. nonnegative real scalars, this polynomial is quadratic. Its expansion into

$$\sum_{k=1}^7 \left((1 - \sigma_k)x_k^2 + \sigma_k(\underline{x}_k + \bar{x}_k)x_k - \sigma_k \underline{x}_k \bar{x}_k \right) \tag{14}$$

shows that its Hessian matrix, $H = \text{diag}((1 - \sigma_1), (1 - \sigma_2), \dots, (1 - \sigma_7))$, is diagonal. This polynomial is then concave if H is negative-definite which happens for $\sigma_k > 1$ for $k = 1 \dots 7$. This shows that N and nonnegative scalars (SoS) $\sigma_k, k = 1 \dots 7$, do exist for $\sigma_0(x)$ to be PSD. Furthermore, σ_k being scalars, $\sigma_0(x)$ is a quadratic polynomial. As discussed in Sect. 2, Hilbert (1888) showed that, for quadratic polynomials, every PSD polynomial is SoS. \square

Let us now consider the problem of checking whether or not (9) is SoS when considering the polynomials $g_k(x), k = 1 \dots 7$ of Proposition 1. If so the assignment a is definitely an outlier within the bounds. If one knows beforehand that $\lambda_i^a f_i^a(x) - \xi$ must be positive, a sequence of $\sigma_k(x)$ of increasing degree can be used until a positivity certificate is obtained. However, for the problem at hand, when a set of $\sigma_k(x)$ of some degree fails to deliver such certificate, it is either because $\lambda_i^a f_i^a(x) - \xi$ is indeed not positive or the required degree for a positivity certificate has not been reached. The good news here is that, within a BnB search, the considered bound intervals $[\underline{x}, \bar{x}]$ get smaller and we show in the following that using nonnegative scalars σ_k rather than SoS polynomials of higher degree suffices. To see this, consider the following proposition:

Proposition 2 *Let $\hat{x} \in \mathbb{R}^7$ with known entries. The following statements are equivalent*

- (i) $\lambda_i^a f_i^a(\hat{x}) - \xi > 0$.
- (ii) *There exist SoS polynomials $\sigma_k(x), k = 1 \dots 7$:*

$$\lambda_i^a f_i^a(x) - \xi + \sum_{k=1}^7 (x_k - \hat{x}_k)^2 \sigma_k(x) > 0. \tag{15}$$

If $f_a(x)$ is a polynomial of degree 2, nonnegative scalars σ_k , instead of $\sigma_k(x)$ for all k , is sufficient.

Proof (ii) \implies (i) is straightforward. For (i) \implies (ii), since $\lambda_i^a f_i^a(x) \in M(g)$, the SoS polynomial of (9) can be expressed as,

$$\begin{aligned} \sigma_0 &= \lambda_i^a f_i^a(x) - \xi - \sum_{k=1}^7 \sigma_k(x) g_k(x), \\ &= \lambda_i^a f_i^a(x) - \xi + \sum_{k=1}^7 (x_k - \hat{x}_k)^2 \sigma_k(x). \end{aligned} \tag{16}$$

When $f_i^a(x)$ is a polynomial of degree 2, let G_f be the Gram matrix of $\lambda_i^a f_i^a(x) - \xi$ and G_x be that of $\sum_{k=1}^7 (x_k - \hat{x}_k)^2$. These matrices are defined by: $\lambda_i^a f_i^a(x) - \xi = \mathbf{Z}_1(x)^T G_f \mathbf{Z}_1(x)$ and $\sum_{k=1}^7 (x_k - \hat{x}_k)^2 = \mathbf{Z}_1(x)^T G_x \mathbf{Z}_1(x)$. Note that G_x is PSD and can be written as $G_x = U^T U$ with $U \hat{x} = 0$. The Gram matrix of the polynomial in (15) is then

written as $G_f + U^T \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_7) U$. A direct application of Lemma 1 is that the latter matrix is positive-definite if and only if $\mathbf{Z}_1(\hat{x})^T G_f \mathbf{Z}_1(\hat{x}) > 0$. This, not only shows (i) \implies (ii), but also provides their equivalence. \square

Although, the equivalence can also be proven using one σ instead of multiple σ_k for the case of degree 2 polynomials, the Proposition 2 is only a proof that the outlier certificate can definitely be obtained when the bounding intervals get smaller during the BnB search. Since we are interested in detecting outliers as early as possible, we suggest using multiple σ_k .

We now state the following preliminary result:

Result 3 (Preliminary) *Let $\xi \geq 0$ be a predefined threshold distinguishing inliers from outliers. Consider two vectors \underline{x} and \bar{x} in \mathbb{R}^7 whose respective entries \underline{x}_k and \bar{x}_k satisfy $\underline{x}_k \leq \bar{x}_k$ for $k = 1 \dots 7$. Let \mathcal{K}_b be the set*

$$\mathcal{K}_b = \{x \in \mathbb{R}^7 : g_k(x) := (x_k - \underline{x}_k)(\bar{x}_k - x_k) \geq 0\}, \tag{17}$$

$a \in \mathcal{A}$ be either a point-to-plane or plane-to-plane putative assignment, and $f^a(x)$ the vector of polynomials $f_i^a(x), i = 1 \dots p$, induced by this assignment [(4) for point-to-plane and (7) for plane-to-plane].

If, for at least one of the polynomials $f_i^a(x)$, a scalar $-1 \leq \lambda_i^a \leq +1$ and nonnegative scalars σ_k exist such that

$$\lambda_i^a f_i^a(x) - \xi - \sum_{k=1}^7 g_k(x) \sigma_k \tag{18}$$

is SoS, then $\lambda_i^a f_i^a(x) > \xi$, hence $|f_i^a(x)| > \xi$, for every $\underline{x}_k \leq x_k \leq \bar{x}_k$. In this case, assignment a is then guaranteed to be an outlier (a point-to-plane or plane-to-plane mismatch) within these bounds. Assignment a is a potential inlier only if (18) is not SoS for any $i = 1 \dots 4$. Note that, in the case of a plane-to-plane assignment a , the latter is deemed an outlier only if (18) is found to be SoS for both values ± 1 of δ in (7).

Furthermore, a consequence of Proposition 2 is that when $\bar{x}_k - \underline{x}_k$ tends towards zero, we are guaranteed that any outlier within the bound is detected. Indeed, this can be seen by noticing that when $\bar{x}_k = \underline{x}_k = \hat{x}_k$, polynomial (18) turns into (15).

Whether (18) is SoS can be tested by converting it into its corresponding Gram matrix LMI feasibility problem for the λ_i^a and σ_k indeterminates. Both λ_i^a and σ_k appear in their respective Gram matrices while searching for the SoS certificate of (18). Although the guarantee of identifying outliers (using scalar multipliers σ_k) is demonstrated with a zero-gap bound, in practice, outliers are detected very early in the process. As demonstrated in our experiments, the ability to detect outliers is improved with every size reduction of the investigated bounds. It may be tempting to use higher degree

$\sigma_k(x)$ SoS polynomials to boost the process. However, this is unnecessary and yields slower performances compared to branching.

6 Semi-Algebraic Sets for Geometric Constraints

Recall that our goal is to register a SfM-induced reconstruction of points or planes and a plane-segmented scanned scene. Unlike when dealing with 3D-3D registration, additional geometric constraints emanating from the cameras can be exploited. Some may be implicit, such as plane visibility, others, such as vague camera locations, may be obtained from extra knowledge. In addition, when dealing with segmented scanned scenes, one is given planar patches rather than infinite planes. These patches can be used for point-to-patch in lieu of point-to-plane registration to restrict the location of the SfM-induced points patches. In all cases, these constraints are described by quadratic polynomial inequalities that may augment the semi-algebraic set \mathcal{K}_b derived from the bound constraints. Indeed, adding new polynomial inequalities to those in \mathcal{K}_b has no effect on the Archimedean property of its quadratic module and Proposition 2 still holds. Note though that some of the constraints presented herein, which we refer to as “generic constraints”, may be exploited in both point-to-plane and plane-to-plane registration. Others are applicable only to point-to-plane registration.

6.1 Generic Constraints

Whether engaging in point-to-plane or plane-to-plane, extra-knowledge about vague camera locations or scale information can be used as additional registration constraints.

Camera Bounds A camera center C may lie within a box delimited by six planes in the set $\Psi = \{\psi_k\}_{k=1}^6$ defined by their normal vectors ψ_k and signed distances d_k , as illustrated in Fig. 3(left). Such information can be obtained from application-specific knowledge (GPS, moving vehicle, etc.). This knowledge can be used for further enforcing the search for point-to-plane or plane-to-plane outliers and turns very useful when no putative correspondences are initially known. Consider the cartesian coordinate vector c of the camera center and let

$$\mathcal{K}_c = \{x \in \mathbb{R}^7 : h_k(x) := (\psi_k^T(Qc + t) - d_k)\delta_k \geq 0, \quad k = 1 \dots 6\} \tag{19}$$

where δ_k is the known sign, with respect to Φ_k , of any point within the considered box. If $h_k(x)$ are positive, the camera center is within the box. One can now test if $\lambda_i^a f_i^a(x) - \xi > 0$ whenever the camera center is in the box defined by \mathcal{K}_c .

Quaternions and Scale In the absence of scale, quaternion parameters demand that $q^T q = 1$. When dealing with a scaled scene, the rotation is represented by a scaled quaternion matrix and one can only enforce that $q^T q > 0$. It is understood that, in order to keep the problem numerically tractable via the Archimedean property, all registration parameters need to be bounded. The scale of the scene is no exception. When a better lower bound $\underline{s} > 0$ on the scale s is available, it is advised to enforce that $q^T q \geq \underline{s}$. This condition does not appear in the set \mathcal{K}_b and hence must be accounted for. Assuming the entries $x_k, k = 1 \dots 4$ of x correspond the quaternion parameters, we consider the set

$$\mathcal{K}_q = \left\{ x \in \mathbb{R}^7 : q(x) := -\underline{s} + \sum_{k=1}^4 x_k^2 \geq 0 \right\} \tag{20}$$

Furthermore, since both q and $-q$ yield the same rotation matrix, the initial lower bound of one of the quaternion parameters may arbitrarily be chosen nonnegative. The rest of the quaternion parameters may be initially bounded between $-\sqrt{\bar{s}}$ and $\sqrt{\bar{s}}$ where \bar{s} is the upper bound of the scale.

6.2 Additional Point-to-Plane Geometric Constraints

In addition to camera bounds and scale constraints, other constraints can be used when dealing specifically with point-to-plane registration.

Patches Consider a plane Π , from the scanned scene, and three or more planes Φ_k , not necessarily from the scene, orthogonal to it. The Φ_k planes must be chosen such that their intersection with Π defines a convex region on Π . The set of points on Π within this convex region is a patch. In practice, four such planes are adequate to represent meaningful patches in man-made environments, as shown in Fig. 3(middle). Each Φ_k is described by its normal vector ϕ_k and signed distance d_k . Let us denote by Φ the set $\{\Phi_k\}_{k=1}^4$ and let $\delta_k = \pm 1$ be the known sign, with respect to Φ_k , of a scanned point lying within the considered region. In the point-to-patch case, we can then identify outliers by checking whether point-to-plane $f_i^a(x)$ of (4) is positive everywhere within the bounds of x and in the set

$$\mathcal{K}_a^\Phi = \{x \in \mathbb{R}^7 : p_k(x) := (\phi_k^T(Qy + t) - d_k)\delta_k \geq 0, \quad k = 1 \dots 4\} \tag{21}$$

The polynomials in this set indicate the sign of the point Y , with coordinates y , with respect to each Φ_k .

Plane Visibility Consider a point Y on a scene plane Π . If this point is imaged by two cameras, then these can only observe the same side of the plane: the one on which the point lies, as shown in Fig. 3(right). In order for the cameras to observe the same side of the plane, their camera centers must lie on

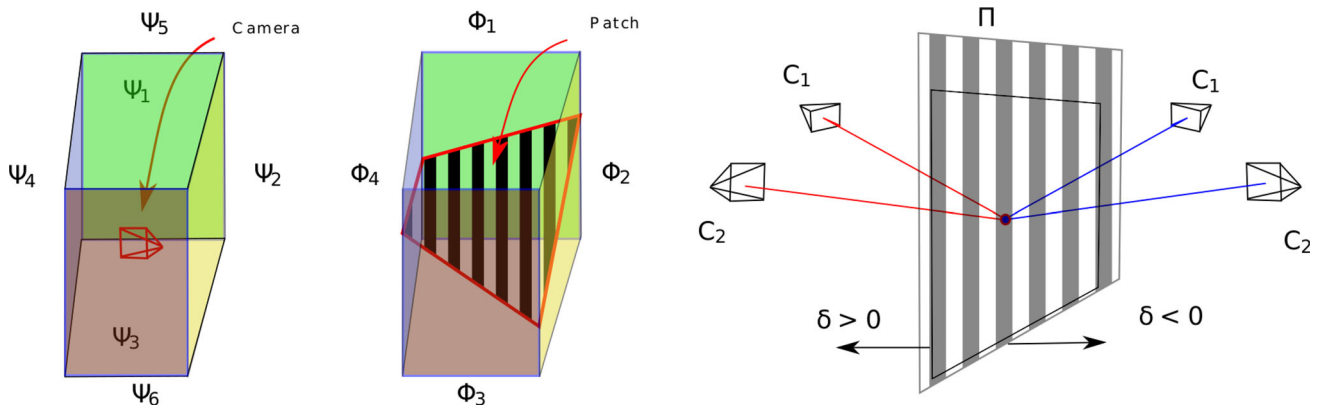


Fig. 3 Illustrations of camera bounds delimited by six planes (left) defining a bounding box, plane as a patch delimited by four planes (middle), and plane visibility for two sets of cameras lying on two different sides (right) (Color figure online)

one side with respect to Π . Camera centers can easily be obtained from the SfM-calculated camera matrices: they are their right null space. Let C_k be the camera centers of $n \geq 2$ cameras with cartesian coordinates c_k . We define the set $\mathcal{K}_{\Pi}^{\delta}$ such that

$$\mathcal{K}_{\Pi}^{\delta} = \{x \in \mathbb{R}^7 : v_k(x) := (\pi^T(Qc_k + t) - d)\delta \geq 0, \\ k = 1 \dots n\} \quad (22)$$

where $\delta = \pm 1$. We denote \mathcal{K}_{Π}^{+} the set $\mathcal{K}_{\Pi}^{\delta}$ obtained using $\delta = +1$ and \mathcal{K}_{Π}^{-} otherwise. A given assignment a is a definite outlier if $f_a(x) > 0$ in \mathcal{K}_{Π}^{+} and in \mathcal{K}_{Π}^{-} (in addition to patch and bounds conditions). Furthermore, planes for which $v_1(x)$ and $v_2(x)$ (for two cameras 1 and 2) always have opposite signs within the bounds of x cannot be assigned any points visible in those cameras. This would indicate that the plane always cuts the base-line of the two camera and cannot contain points visible in both cameras. Testing this can be carried out by checking, for $\delta = \pm 1$, whether

$$\begin{cases} \exists \sigma_k : v_1(x) - \sum_{k=1}^7 g_k(x)\sigma_k \text{ is SoS,} \\ \exists \sigma_k : -v_2(x) - \sum_{k=1}^7 g_k(x)\sigma_k \text{ is SoS.} \end{cases} \quad (23)$$

If for both values of δ , each polynomial in (23) is SoS, plane Π shall not be considered for assigning SfM points emanating from those cameras.

7 Registration

The preliminary results in Result 3 apply to both point-to-patch and plane-to-plane registration problems. Based on these results and the semi-algebraic sets presented in Sect. 6, we are now ready to state additional results for point-to-plane, or rather plane-to-patch, and plane-to-plane registration. These results along with those in Result 3 are

used in our BnB registration algorithm which is also presented in this Section.

Result 4 (Point-to-patch) *Let $\xi \geq 0$ be a predefined threshold distinguishing inliers from outliers. Assume we are given a putative point-to-plane assignment $a = (Y, \Pi) \in \mathcal{A}$, a patch on Π delimited by the planes in the set $\Phi = \{\Phi_k\}_{k=1}^4$, lower \underline{x} and upper \bar{x} bounds on the registration parameter vector x , bounds \underline{s} and \bar{s} on the scale of the scene, and (optionally) bounds defined by planes $\Psi = \{\Psi_k\}_{k=1}^6$ on the location of the camera centers of one (possibly more) camera. One would like to know whether or not the SfM-reconstructed point Y may lie on Π , while Π is visible by the cameras observing Y , within the patch Φ with registration parameters in the bounds \underline{x} and \bar{x} . In order to establish whether such assignment is possible, we consider the set*

$$\mathcal{K} = \{x \in \mathbb{R}^7 : x \in \mathcal{K}_b \cap \mathcal{K}_a^{\Phi} \cap \mathcal{K}_{\Pi}^{\delta} \cap \mathcal{K}_c \cap \mathcal{K}_q\} \quad (24)$$

resulting from the intersection of all the sets defined by (17), (21), (22), (19) and (20). If there exist a scalar $-1 \leq \lambda^a \leq +1$ and nonnegative scalars $\sigma_k, \sigma'_k, \sigma''_k, \sigma'''_k$ and σ such that [considering $f^a(x)$ given by (4)]

$$\lambda^a f^a(x) - \xi - \sum_{k=1}^7 g_k(x)\sigma_k - \sum_{k=1}^4 p_k(x)\sigma'_k \\ - \sum_{k=1}^n v_k(x)\sigma''_k - \sum_{k=1}^6 h_k(x)\sigma'''_k - q(x)\sigma \quad (25)$$

is SOS, then $|f^a(x)| > \xi$ in \mathcal{K} and the assignment a is a definite outlier. It is a potential inlier otherwise.

Note that a point-to-plane version of this results can simply be obtained by not using \mathcal{K}_a^{Φ} to construct \mathcal{K} .

Result 5 (Plane-to-plane) *Let $\xi \geq 0$ be a predefined threshold distinguishing inliers from outliers. Assume we are given a putative plane-to-plane assignment $a = (\Pi_r, \Pi) \in \mathcal{A}$, lower \underline{x} and upper \bar{x} bounds on the registration parameter vector x , bounds \underline{s} and \bar{s} on the scale of the scene, and (optionally) bounds defined by planes $\Psi = \{\Psi_k\}_{k=1}^6$ on the location*

of the camera centers of one (possibly more) camera. One would like to know whether or not the SfM-reconstructed plane Π_r may be aligned with Π for registration parameters in the bounds \underline{x} and \bar{x} . In order to establish whether such assignment is possible, we consider the set

$$\mathcal{K} = \{x \in \mathbb{R}^7 : x \in \mathcal{K}_b \cap \mathcal{K}_c \cap \mathcal{K}_q\} \tag{26}$$

resulting from the intersection of all the sets defined by (17), (19) and (20). If, for at least one polynomial $f_i^a(x)$, $i = 1 \dots 4$, there exists a scalar $-1 \leq \lambda_i^a \leq +1$ and nonnegative scalars σ_k, σ'_k and σ such that

$$\lambda_i^a f_i^a(x) - \xi - \sum_{k=1}^7 g_k(x)\sigma_k - \sum_{k=1}^6 h_k(x)\sigma'_k - q(x)\sigma \tag{27}$$

is SOS, then $|f_i^a(x)| > \xi$ in \mathcal{K} and the assignment a is a definite outlier. It is a potential inlier only if (27) is not SoS for any $i = 1 \dots 4$.

Recall that SoS problems (25) and (27) in these results can be solved as a LMI feasibility problem.

Our registration approach is based on Results 4 and 5. In the following, we use the term point-to-plane to refer to both point-to-plane and point-to-patch assignments. Since the goal of the BnB algorithm is to estimate the registration parameters yielding the largest number of inliers, our algorithm is provided either a set of putative point-to-plane or plane-to-plane correspondences. In the absence of such correspondences, we consider every SfM-induced point or plane to be putatively assigned to all the planes (or patches) in the scanned scene. A dynamically-built search tree, whose nodes are registration parameter bounds, allows to explore the space of parameters, as discussed in Sect. 3.

Optimistic Inliers Given a set of putative assignments and the semi-algebraic set for parameter bounds, our algorithm estimates the optimistic number of potential inliers $\bar{\eta}$ (and the corresponding inlier set $\bar{\mathcal{I}}$) using Results 4 and 5, with appropriate additional semi-algebraic sets for generic constraints. To qualify an assignment to be an inlier, we distinguish two cases:

1. Putative point-to-plane (resp. plane-to-plane) correspondences are provided: a point-to-plane (resp. plane-to-plane) assignment qualifies as a potential inlier if (25) (resp. 27) is not proven SoS for any polynomial in vector $f^a(x)$ for the given assignment.
2. No putative correspondences are provided: all possible point-to-plane (resp. plane-to-plane) assignments are tested. An assignment is considered a potential inlier if (25) (resp. 27) is not proven SoS for any polynomial in vector $f^a(x)$ of that assignment.

Now we summarize the process of estimating optimistic number inliers in Algorithm 2.

Algorithm 2 Optimistic Inliers Estimation

Input: \mathcal{K}, \mathcal{A} **Output:** $\bar{\eta}, \bar{\mathcal{I}}$
 Initialization: $\bar{\mathcal{I}} = \mathcal{A}$
 1: **for** each $a \in \mathcal{A}$ **do** ▷ Using (25) or (27)
 2: **if** $\exists i \in \{1, \dots, p\}$:
 3: $\exists \lambda_i^a \in [-1, +1], \lambda_i^a f_i^a(x) > \xi, \forall x \in \mathcal{K}$ **then**
 4: $\bar{\mathcal{I}} = \bar{\mathcal{I}} - a$
 5: **end if**
 6: **end for**
 7: $\bar{\eta} = |\bar{\mathcal{I}}|$ ▷ Number of inliers

Pessimistic Inliers A local refinement method is used to obtain a pessimistic number of inliers $\underline{\eta}$ (and the corresponding inlier set $\underline{\mathcal{I}}$) for each node. To maximize the pessimistic number of inliers, a local method iteratively refines the registration parameters. The refinement process starts from the mid-values of the registration parameter bounds. In order to be representative of the node, it searches the optimal solution within the investigated bounds. Given the set of optimistic inlier assignments $\bar{\mathcal{I}}$ (obtained from Algorithm 2) and the semi-algebraic set \mathcal{K}_b representing the bounds under consideration, the algorithm iteratively updates the registration parameters to:

$$\hat{x} = \operatorname{argmin}_{x \in \mathcal{K}_b} \sum_{a \in \bar{\mathcal{I}}} \|f^a(x)\|^2. \tag{28}$$

The process of estimating the pessimistic number inliers is given by Algorithm 3.

Algorithm 3 Pessimistic Inliers Estimation

Input: $\mathcal{K}_b, \bar{\mathcal{I}}$ **Output:** $\underline{\eta}, \underline{\mathcal{I}}$
 1: Obtain \hat{x} for \mathcal{K}_b and $\bar{\mathcal{I}}$.
 Initialization: $\underline{\mathcal{I}} = \emptyset$
 2: **for** each $a \in \bar{\mathcal{I}}$ **do**
 3: **if** $\forall i \in \{1, \dots, p\}, |f_i^a(\hat{x})| \leq \xi$ **then**
 4: $\underline{\mathcal{I}} = \underline{\mathcal{I}} \cup a$
 5: **end if**
 6: **end for**
 7: $\underline{\eta} = |\underline{\mathcal{I}}|$ ▷ Number of inliers

To summarize, the optimistic and pessimistic inliers are obtained using Algorithm 2 and Algorithm 3, respectively. Then, the registration parameters yielding the largest number of inliers are obtained using Algorithm 1. In our implementation, the qualified node is branched along its longest edge resulting in two new nodes to be processed. The node corresponding to the maximum number of pessimistic inliers is processed first.

8 Discussion

In general, our method converges while the explored bounds are still quite large. This suggests that incorporating ξ in the SoS test plays a limited role and most of the time can be ignored. What matters most is that potential inliers, as per the local method's measure, not to be considered outliers by the SoS test. Without the threshold ξ , the SoS test is of the form (dropping λ for the sake of the argument) $f(x) - \sum_{k=1}^7 \sigma_k(x_k - \underline{x}_k)(\bar{x}_k - x_k) > 0$ which, when true, means $f(x) > \sum_{k=1}^7 \sigma_k(x_k - \underline{x}_k)(\bar{x}_k - x_k)$ for all x within bounds.

No assignment would be mistakenly considered as an outlier by the SoS test so long as $\sum_{k=1}^7 \sigma_k(x_k - \underline{x}_k)(\bar{x}_k - x_k) > \xi$ for all x within the bounds. Note that $\sum_{k=1}^7 \sigma_k(x_k - \underline{x}_k)(\bar{x}_k - x_k)$ is concave and its maximum is $\frac{1}{4} \sum_{k=1}^7 \sigma_k(\bar{x}_k - \underline{x}_k)^2$. Therefore, the SoS test not considering ξ would be safe as long as it produces no σ_k such that $\frac{1}{4} \sum_{k=1}^7 \sigma_k(\bar{x}_k - \underline{x}_k)^2 < \xi$.

Note that σ_k compensate the nonconvexities in $f(x)$ and are generally relatively large. Hence for the SoS test to be unsafe when ξ is not included, the size of the bound-gap $(\bar{x}_k - \underline{x}_k)^2$ for $k = 1 \dots 7$ must be very small. The algorithm converges before reaching such small bound.

9 Experiments

We conducted experiments with seven different benchmark real datasets shown in Fig. 4 whose details can be found in Jensen et al. (2014) and Strecha et al. (2008). Both point-to-plane (more precisely, point-to-patch) and plane-to-plane registration methods have been tested. Our algorithm was implemented in MATLAB2014b and the SoS problems were solved using the LMI Control Toolbox.¹ All experiments were carried out on a 8GB RAM Pentium i7/3.40GHz. The SfM reconstructions and segmented scene planes were obtained using the openMVG Toolbox (Moulon et al. 2013) and Hough Transform based plane detector (Borrmann et al. 2011). All the planes were extracted directly from the point clouds using the source code provided in Hough Transform Plane Detector (2015) with an accumulator designed in Borrmann et al. (2011) to speed up the Hough Transform. For the experiments with unknown scale of the reconstruction, the initial bound on reconstruction scale was set to 0.2–5.0 (five times in scale in both directions). Four different error measurement metrics were defined to evaluate the registration quality: the RMS 3D error on normalized point sets, errors in rotation R , translation \mathbf{t} , and scale s . For N experiments,

these are defined as follows :

$$\Delta R = \sqrt{\frac{1}{3N} \sum_{i=1}^N \|r_i^* - r\|^2}, \Delta T = \sqrt{\frac{1}{N(\|\mathbf{t}\|^2)} \sum_{i=1}^N \|\mathbf{t}_i^* - \mathbf{t}\|^2},$$

$$\Delta S = \sqrt{\frac{1}{N(s^2)} \sum_{i=1}^N (s_i^* - s)^2},$$

where r is a vector obtained by stacking three rotation angles in degrees. The estimated variables are represented with * and variables without it are their ground truth.

Note that all SoS inlier/outlier tests have been carried out with $\xi = 0$.

9.1 Point-to-Plane Registration

To perform point-to-plane registration, we selected only the prominent planes from the Euclidean scene. This allowed us to select more points while keeping the exhaustive combinations between points and planes within practical limits. It is important because, when a set of sparse points are selected randomly, we wish to ensure that at least 7 points lying on 4 different planes are included. In this case, we conducted two sets of experiments : one with and one without putative point-to-plane correspondences.

9.1.1 Inlier Set Maximization with Correspondences

The method was first tested for known putative correspondences where the synthetic inliers/outliers were generated under real data setups. No bounds on cameras were used in these experiments. To test robustness to outliers, we varied the number of outliers up to 90% for Scene73 and compared the results against the linear 12-point RANSAC. In theory, only 7 point-to-plane assignments (at least of 4 different planes) are sufficient to estimate the similarity transformation. Therefore, 7-point RANSAC is preferred over 12-point RANSAC, if an efficient minimal solver is available for 7 point-to-plane assignments. Up to our knowledge, there exists no such minimal solver. Ideally, generic polynomial solvers may also serve the purpose. But unfortunately, our attempts with such solvers for minimal correspondences case (7 equations, 7 variables, and 2 degree, form 7 point-to-plane correspondences), computation time was significant enough to render this solution impractical within the framework of RANSAC.

Figure 5 shows that our method consistently detects 21 inliers for every experiment while RANSAC fails to detect the least number of required inliers starting from 45% of outliers. Note that the numbers of inliers reported here are true positive inliers. Furthermore, our method does not detect any false positive inliers. Figure 6 shows the errors in rotation,

¹ Our source code can be found at: www.dropbox.com/sh/vvkeaf5fcaxwsyr/AACqTUJE3FTXeiPnFkCbOSSXa?dl=0.

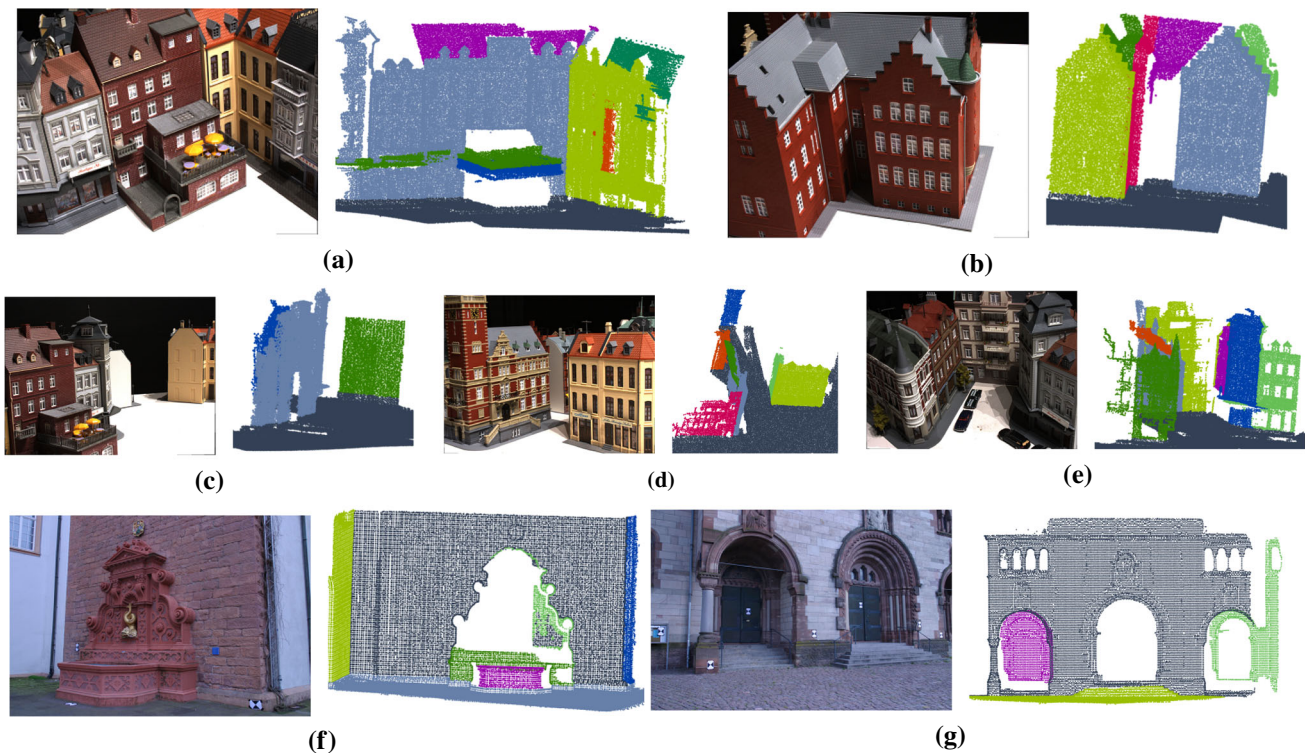


Fig. 4 Sample image and the segmented scene (shown in different colors for each plane) of seven different datasets. **a** Scene23, **b** Scene24, **c** Scene27, **d** Scene29, **e** Scene73, **f** Fountain and **g** Herz-Jesu (Color figure online)

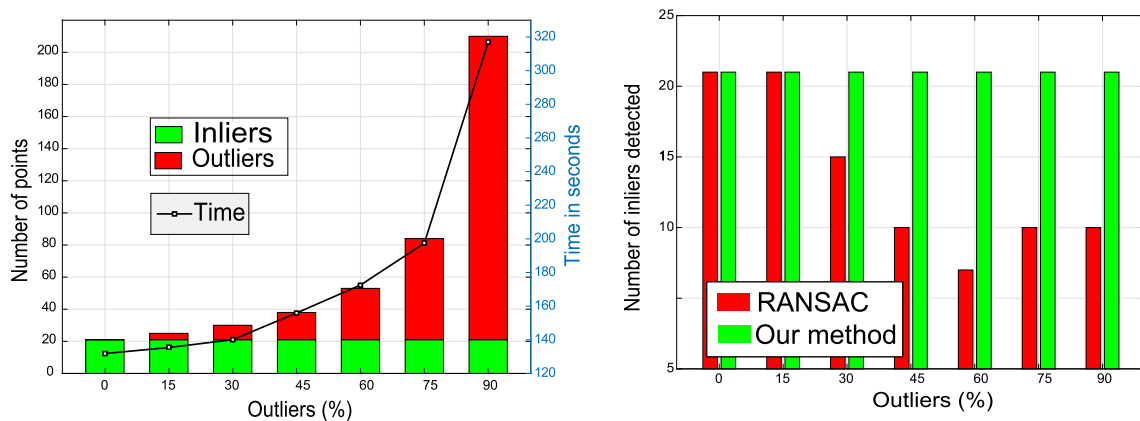


Fig. 5 Point-to-plane registration for Scene73 with correspondences and no camera bounds. Left: number of processed points and time taken for various levels of outliers. Right: number of detected inliers using RANSAC and our method (Color figure online)

translation, and scale for the same scene with various levels of outliers. The convergence graph of our method with 50% outliers is shown in Fig. 7 for Scene23, Scene73, and Fountain whose quantitative results are shown in Table 1. Figure 8 shows the evolution of the volume and the number of nodes remaining to be processed for the first 50 iterations on Scene23 with 50% outliers. The qualitative results for scene73 are shown in Fig. 9.

Furthermore, we conducted experiments with the minimal solver of Ramalingam and Taguchi (2013) within RANSAC

while allowing it to run for as long as the proposed method did. Note that the minimal solver Ramalingam and Taguchi (2013) is designed for rigid body transformations. Therefore, the reconstruction scale needs to be known beforehand. We favored the solver of Ramalingam and Taguchi (2013) by providing the exact ground truth scale. The experiments with Ramalingam and Taguchi (2013) for Scene73 are reported in Fig. 10. Observe that RANSAC with Ramalingam and Taguchi (2013) achieves a success rate of 100% for outliers upto 50%. The success rate then drops to 69% for 60% of

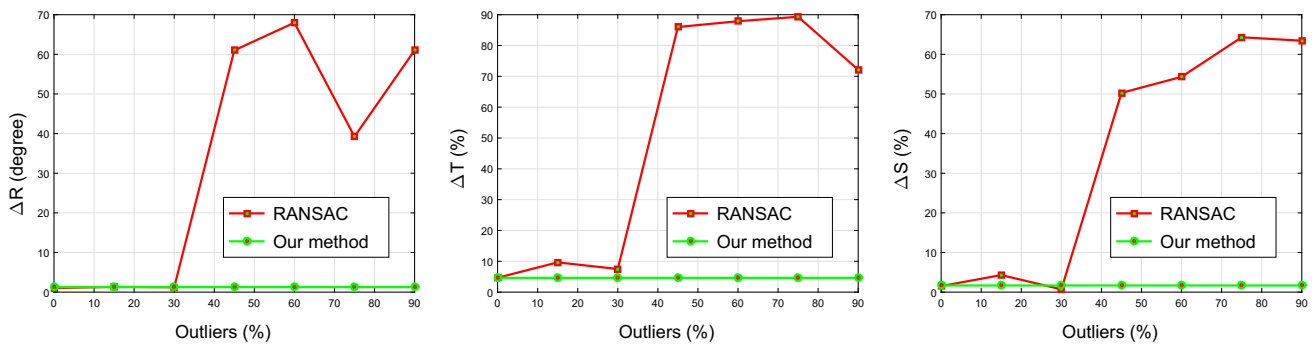


Fig. 6 Point-to-plane registration for Scene73 with increasing number of outliers, with correspondences and no camera bounds. Left: error in rotation. Middle: error in translation. Right: error in scale (Color figure online)

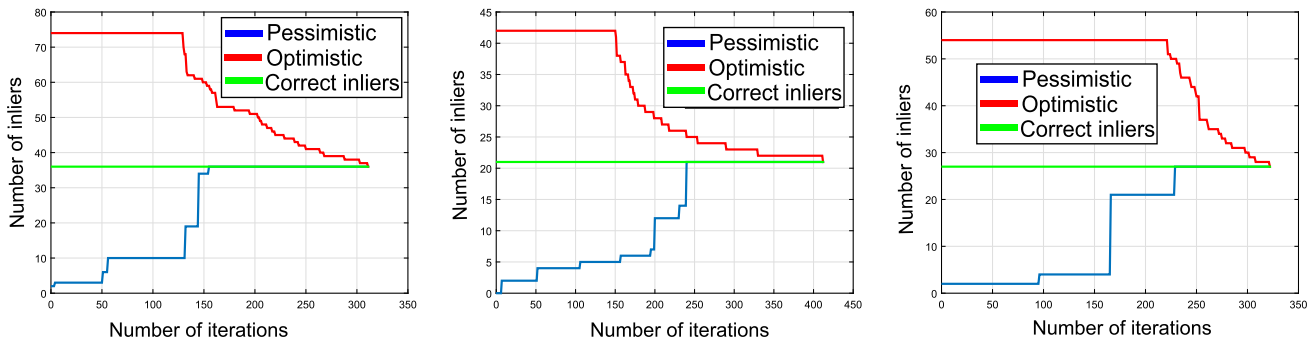


Fig. 7 Point-to-plane registration convergence graph for 50% outliers, with correspondences and no camera bounds. Left: Scene23. Middle: Scene73. Right: Fountain (Color figure online)

Table 1 Point-to-plane registration with correspondences and no camera bounds: quantitative results obtained with 50% outliers

	ΔR (degree)	ΔT (%)	ΔS (%)	3D error	Time (s)
Scene23	0.785	1.75	0.21	0.0163	168.95
Scene73	1.263	4.63	1.68	0.0219	153.39
Fountain	0.524	1.21	0.53	0.0056	546.41

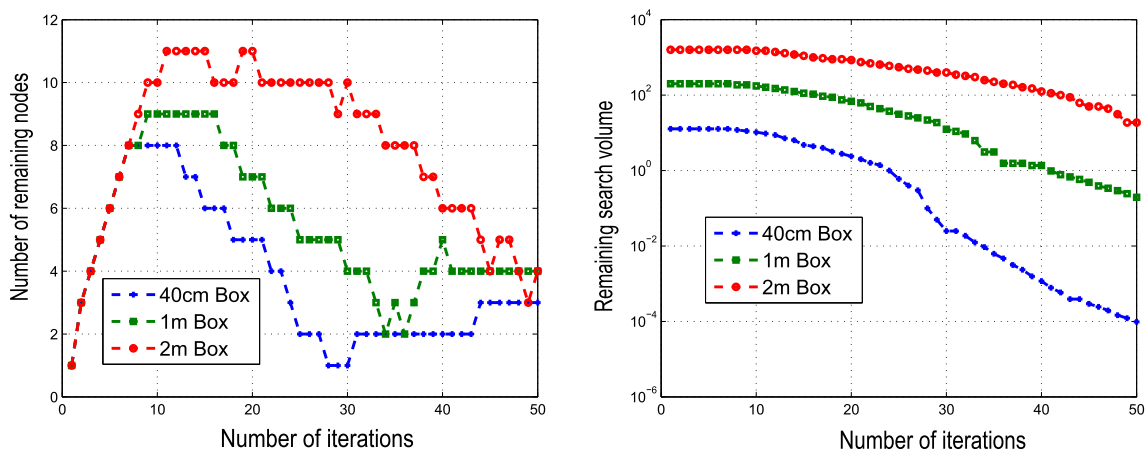


Fig. 8 Point-to-plane registration for Scene23 with correspondences, no camera bounds, and 50% outliers. Left: number of remaining nodes versus search iterations. Right: number of remaining volume versus search iterations (Color figure online)



Fig. 9 Point-to-plane registration for Scene73 with correspondences. Left: recovered point-to-plane correspondences with all the processed points in red. Right: registration between reconstruction (in green) and the segmented scene (Color figure online)

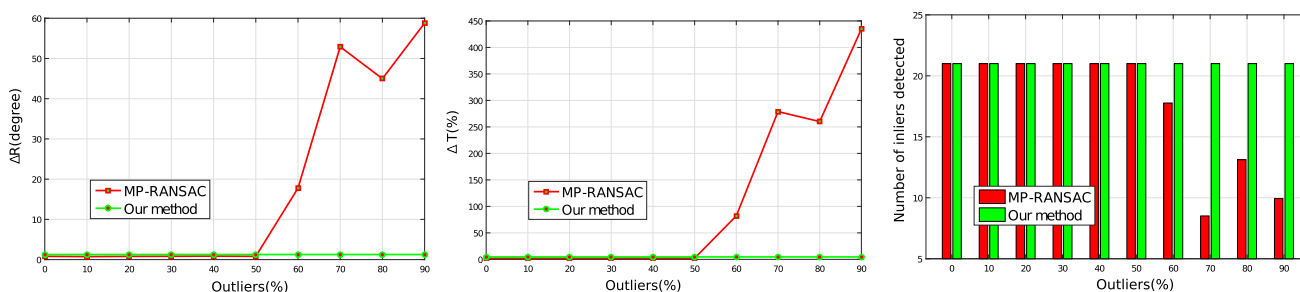


Fig. 10 Left to right: errors in rotation, errors in translation, and number of detected inliers obtained by our method compared against Minimal-Point Ransac with known scale (Ramalingam and Taguchi 2013). Experiments conducted with Scene73 (Color figure online)

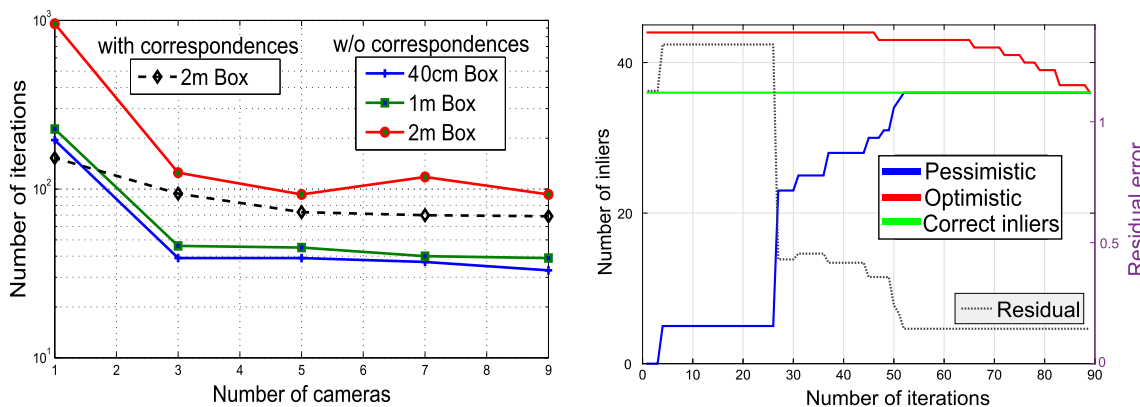


Fig. 11 Point-to-plane registration for Scene23. Left: number of iterations versus number of cameras. Right: Convergence graph for the case w/o correspondences with three 1m-box bounded cameras (Color figure online)

outliers. All the experiments conducted with more than 70% of outliers failed.

9.1.2 Inlier Set Maximization w/o Correspondences

In the absence of initial correspondences, each point was assigned to all available planes. We conducted several experiments with bounded cameras by changing the number of

bounded cameras and camera bounding box size. The number of iterations taken for these configurations are shown in Fig. 11(left) for Scene23. The average time per iteration is 1.15sec. In the same figure, we also provide the number of iterations taken for the “with correspondences” case with 50% outliers and 2m camera bounding boxes. The case of a single bounded camera is equivalent to unbounded cameras but bounded translation: plane visibility criterion

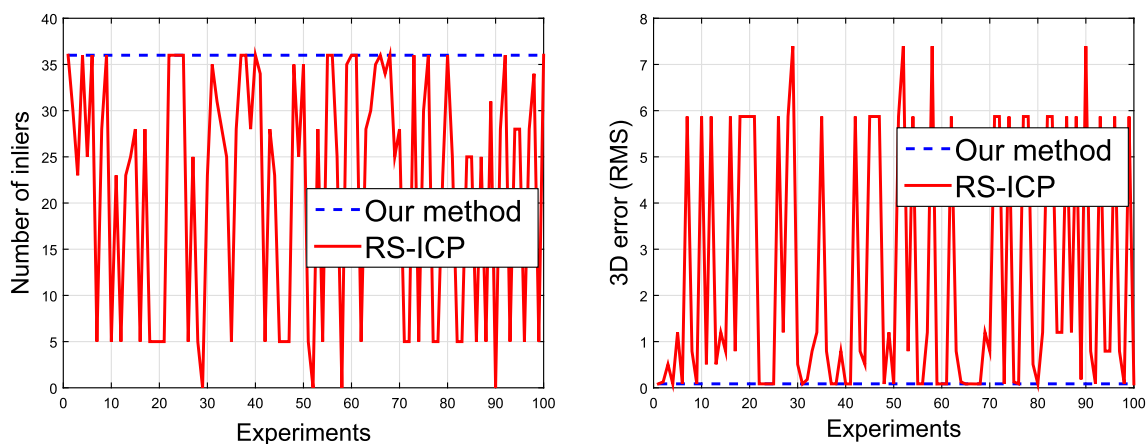


Fig. 12 Point-to-plane registration for Scene23 without correspondences and three 1m-box bounded cameras (100 independent trials). Our method versus randomly started scaled ICP (RS-ICP). Left: no. of inliers detected. Right: 3D RMS error (Color figure online)

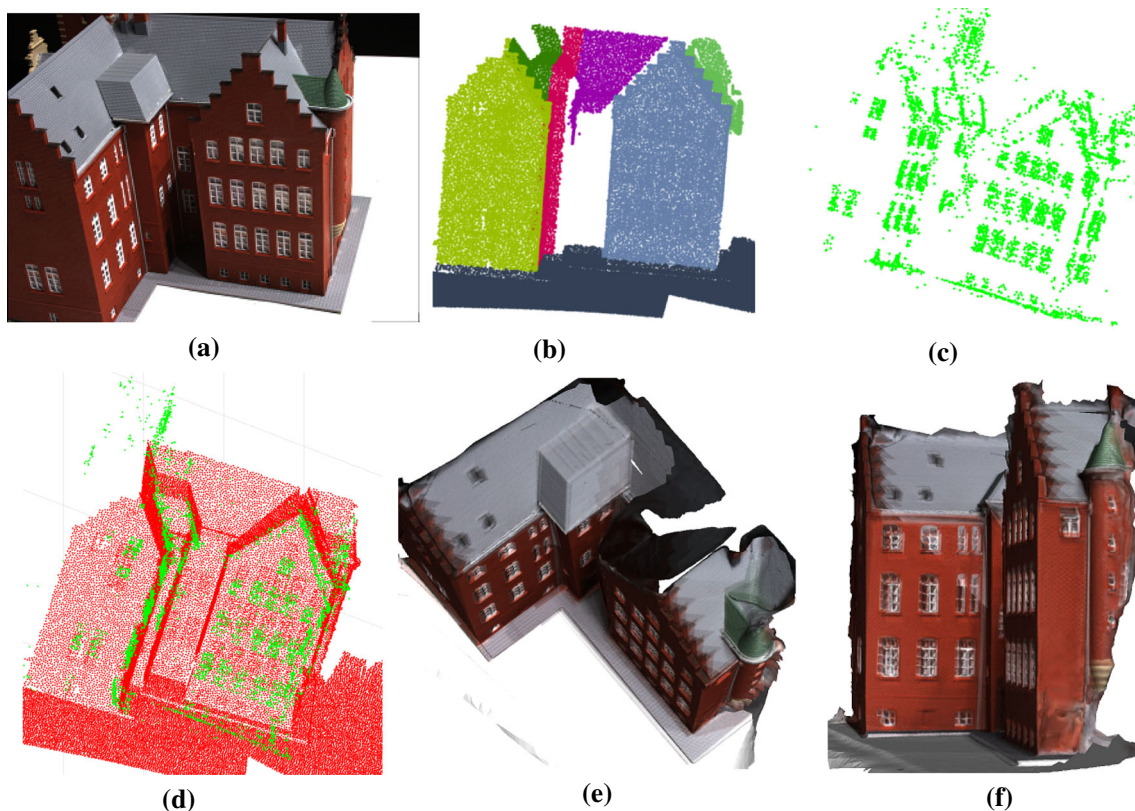


Fig. 13 Point-to-plane registration results for Scene24. **b, c** Inputs, **d** point-to-plane registration and **e, f** texture projection. **a** Sample image, **b** Segmented scene, **c** SfM reconstruction, **d** Registered pointsets, **e** Textured scene, top view and **f** Textured scene, side view (Color figure online)

cannot be used in this case. We recall that initial bounds on all the registration parameters are indispensable to ensure an Archimedean quadratic module of the constraints set and hence for employing Putinar's Psatz. Figure 11(right) shows the convergence graph, using Scene23, obtained with 3 1m-box bounded cameras. It also shows how the residual error on the registration parameters varies with the increase in the number of pessimistic inliers. The reported

box size is for a normalized scene size of about 10 meters. In Fig. 12, we report the results obtained on Scene23 (with 3 1m-box bounded cameras) using our method and a randomly started scaled ICP (RS-ICP) for 100 independent trials. In each trial, the scaled ICP was started at randomly picked registration parameter values satisfying bound and visibility constraints. The results show that, unlike RS-ICP which provides very large 3D errors, our method consis-

Table 2 Point-to-plane registration without correspondences: quantitative results for seven different datasets

Scene	Points	Planes	Recon.	Rep.	Camera	Box	Inlier	ΔR	$\Delta T(\%)$	$\Delta S(\%)$	3D error	Iter	Time (s)
Scene23	90	8	2.08134	0.8373	3	2 m	41	3.4147	1.95	2.48	0.0619	482	599.738
Scene24	47	7	2.62791	0.8756	3	40 cm	31	0.9591	2.31	2.09	0.0424	81	51.572
Scene27	49	4	1.61906	0.8127	8	50 cm	20	2.5759	3.96	1.68	0.0131	133	141.837
Scene29	90	8	1.77408	0.9226	5	1 m	45	2.9995	5.41	2.41	0.081	209	277.351
Scene73	71	8	3.21913	0.8654	5	1 m	45	3.3463	4.78	2.73	0.0654	223	271.226
Fountain	29	7	0.81293	0.8495	4	40 cm	11	2.8639	3.18	4.74	0.0570	102	55.730
Herz-Jesu	129	5	2.08134	0.6402	8	40 cm	101	7.1958	4.02	1.99	0.0464	103	137.766

Table 3 Registration results using RISAG, Go-ICP and our method

Scene	Method	Time (s)	ΔR	$\Delta T(\%)$	3D error
Fountain	RISAG	805.680	8.6825	14.08	0.3275
	Go-ICP	529.415	0.7225	1.63	0.0348
	Our method	55.730	2.8639	3.18	0.0570
Herz-Jesu	RISAG	160.064	17.6378	5.70	0.1830
	Go-ICP	31.254	3.2618	16.9	0.0725
	Our method	137.766	7.1958	4.02	0.0464

tently detects the same number of inliers with the same 3D error. Observe that the success rate of RS-ICP is only about 48%.

The results of our method for all scenes (with their corresponding configurations) are summarized in Table 2. In the reported parameters, Points, Planes, Iter, and Inlier represent their numbers. “Recon.” is the quality of the SfM reconstruction measured as the median reprojection error in pixels while “Rep.” is the fraction of the scene points represented by fitted planes. Observe that the registration quality depends upon the reconstruction quality, representation factor, and the number and size of the camera boxes. For a qualitative evaluation, the results obtained for Scene24 are shown in Fig. 13 along with the registered point sets and textured scene [after further refinement using (Viola and Wells 1997)].

We also provide the results for two datasets obtained using RISAG (Corsini et al. 2013), Go-ICP (Yang et al. 2013), and our method in Table 3. Our method was used without correspondences in the setting given in Table 2. Note that Go-ICP requires an Euclidean reconstruction, which was obtained by upgrading the metric reconstruction using ground truth measurements. Comparison of these methods may be unfair because each requires different initial conditions. Note that the poor performance of RISAG could be due to its RANSAC-driven nature (we used 10^4 RANSAC iterations). Nevertheless, experiments with both RISAG and Go-ICP were conducted in their favorable conditions.

9.2 Plane-to-Plane Registration

In this section, we report the results that we have obtained using our plane-to-plane registration method. The method was tested without any putative correspondences provided, i.e. plane-to-plane registration was carried out on the correspondences obtained by assigning each reconstructed plane to all available scene planes. Experiments with unknown scale as well as without scale ambiguity were carried out. Furthermore, no camera was bounded for all the experiments. The number of planes used for all the datasets are shown in Table 4, where “num. Π_r ” and “num. Π_e ”, respectively, represent the number of reconstructed planes and scene planes.

9.2.1 Inlier Set Maximization with Unknown Scale

The quantitative results obtained for all seven datasets using plane-to-plane registration with scale ambiguity are shown in Table 4. Note that, despite a small number of reconstructed and scene planes, our algorithm was still able to successfully register the data. As expected, the number of search iterations in this case are relatively high. This is due to the absence of constraints from camera bounds. Despite a high number of iterations, the processing time remains reasonable. This is mainly due to the fact that the scenes are represented by a small number of planes leading to a limited number of exhaustive correspondences and hence a reduced number of SoS tests in each iteration.

Convergence graphs for all seven datasets are shown in Fig. 14. For these experiments, the remaining nodes and vol-

Table 4 Plane-to-plane registration results, unknown reconstruction scale, and no camera bounds

Scene	num. Π_r	num. Π_e	Rep.	Inlier	ΔR	ΔT (%)	ΔS (%)	3D error	Iter	Time (s)
Scene23	7	9	0.8736	4	0.8742	9.57	1.95	0.0350	552	486.1561
Scene24	6	10	0.9720	5	2.1486	6.63	3.20	0.0210	458	284.0396
Scene27	6	6	1.00	4	1.5867	3.53	1.23	0.0106	648	413.3966
Scene29	6	9	0.9372	6	1.5340	4.61	1.62	0.0195	437	370.6194
Scene73	9	8	0.8654	7	0.6335	3.00	0.86	0.0227	431	236.7101
Fountain	9	10	0.9824	5	0.3625	8.06	4.09	0.0534	411	381.1078
Herz-Jesu	5	9	0.89.86	5	2.674	11.38	4.35	0.0590	246	124.3724

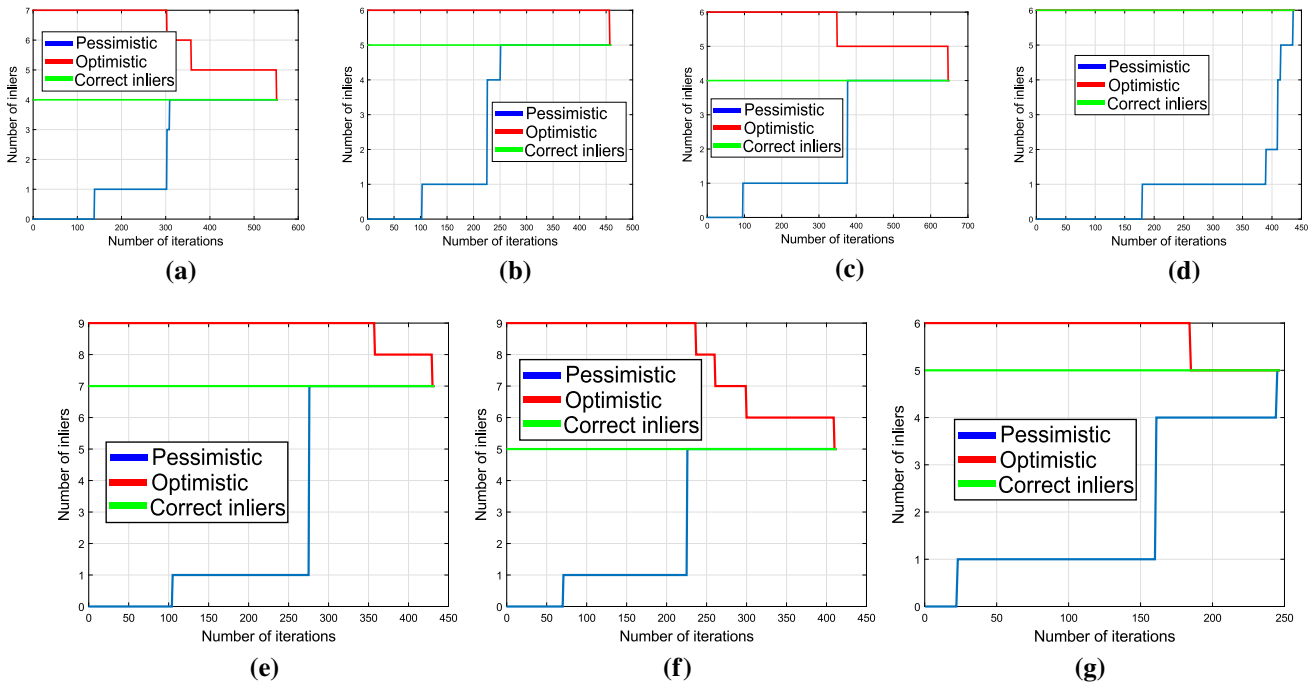


Fig. 14 Plane-to-plane registration convergence graphs for unknown reconstruction scale and no camera bounds. **a** Scene23, **b** Scene24, **c** Scene27, **d** Scene29, **e** Scene73, **f** Fountain and **g** Herz-Jesu (Color figure online)

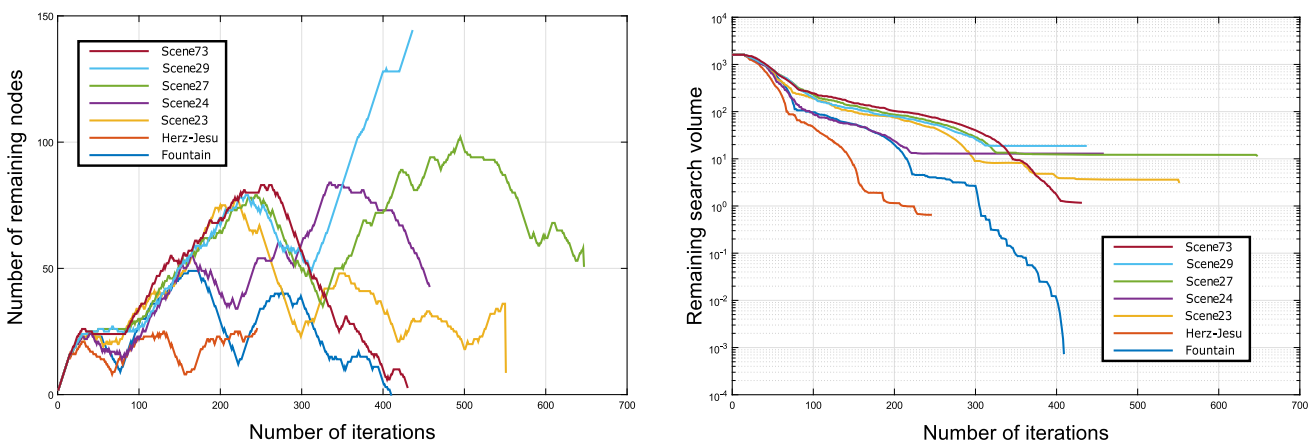


Fig. 15 Plane-to-plane registration with search iterations. Left: remaining nodes. Right: remaining volume (Color figure online)

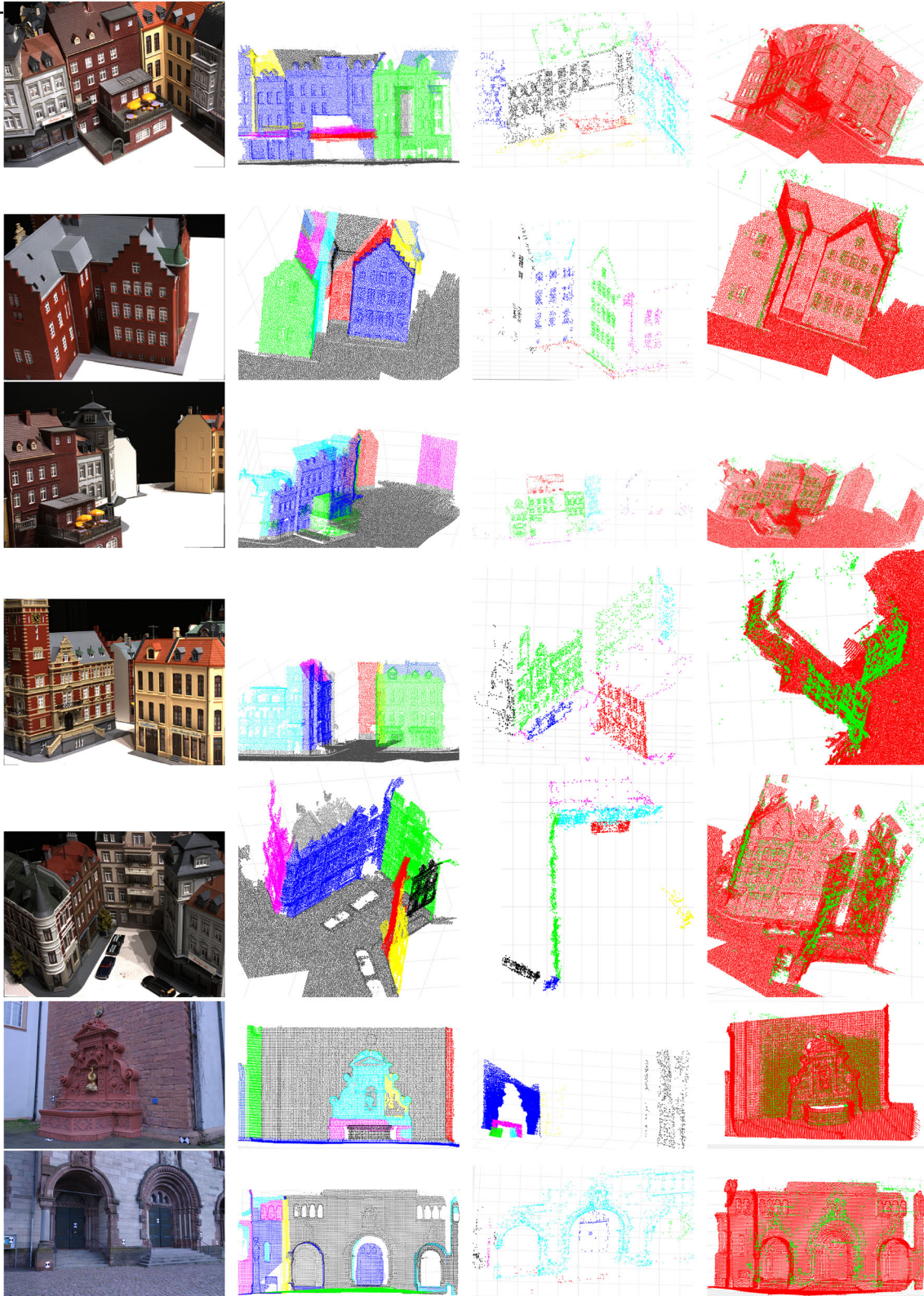


Fig. 16 Left to right: sample image, segmented scene, segmented reconstruction, and plane-to-plane-based registered point sets. (From top to bottom: Scene23, Scene24, Scene27, Scene29, Scene73, Fountain, and Herz-Jesu) (Color figure online)

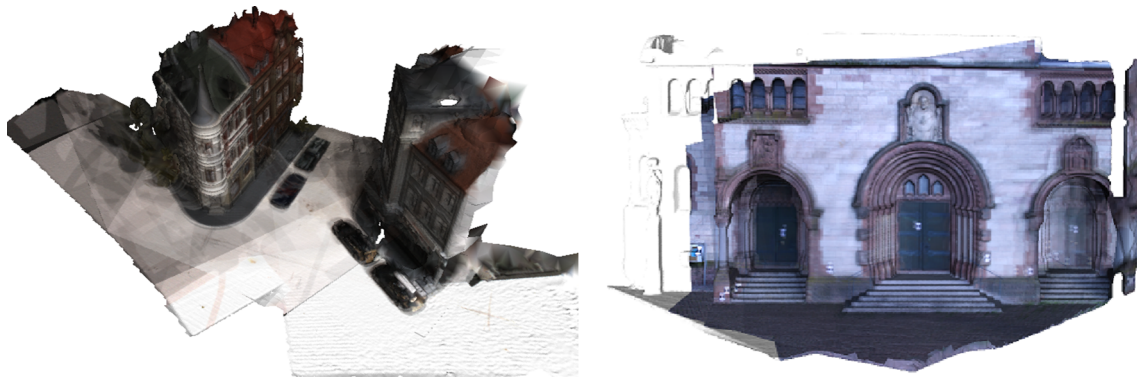


Fig. 17 Texture mapping using plane-to-plane registration results. Left: Scene73. Right: Herz-Jesu (Color figure online)

Table 5 Plane-to-plane registration results, known reconstruction scale, and no camera bounds

Scene	Inlier	ΔR	$\Delta T(\%)$	3D error	Iter	Time (s)
Scene23	4	4.2523	7.57	0.0568	46	48.2516
Scene24	5	2.6900	9.65	0.0571	57	30.0874
Scene27	4	1.5830	4.97	0.0093	72	57.3971
Scene29	6	0.4898	2.58	0.0171	102	99.4266
Scene73	7	0.9946	2.80	0.0206	36	45.7199
Fountain	4	1.3836	8.97	0.0139	33	53.4855
Herz-Jesu	4	0.8295	5.85	0.0982	90	77.7494

ume over the search iterations are shown in Fig. 15. It can be observed that, when the nodes are not efficiently pruned (as in for Scene29 after 300 iterations), the remaining search volume varies only slightly. However, as soon as the maximum number of pessimistic inliers meets that of the optimistic one, the optimal solution is obtained. Thereafter, no remaining nodes/volume needs to be processed.

The qualitative results of plane-to-plane registration with unknown scale and no camera bounds are shown in Fig. 16. Among them, the texture mapping of two datasets, namely Scene73 and Herz-Jesu, obtained after refinement using mutual information maximization, are shown in Fig. 17.

9.2.2 Inlier Set Maximization with Known Scale

In many applications, two point sets are related by a rigid-body transformation. These cases involve no discrepancy of scale between the two point sets to be registered i.e. $\|q\| = 1$. In fact, this registration problem is equivalent to that of 3D-3D registration of structured same-scale point clouds that can both be segmented into planes. The scale information provides two advantages: (1) it reduces the initial bounds on the entries of q ; (2) it introduces an additional polynomial $\|q\|^2 - 1 = 0$ that must always be satisfied for the solution bounds. In our experiments, the reconstruction scales were set to their ground truth values.

The results obtained for known scale plane-to-plane registration are reported in Table 5. As expected, the number of search iterations are significantly smaller than in the case of unknown scale (refer to Table 4). Therefore, the registration process is also faster. Although the same datasets were used for both the cases of unknown and known scale, the sets of inlier planes were found to be different in some cases. This is because of the scale factor relaxation, which was restricted in the latter.

10 Conclusion

We proposed a method for registering a 3D scan and a set of images of a structured scene. The proposed approach is based on the theory of polynomial SoS optimization. Our method uses SoS registration conditions for point-to-plane as well as plane-to-plane registration. The method presented in this paper can incorporate various constraints emanating from scene and camera knowledge (patch segmentation, camera locations, plane visibility, scaling, etc.). Using Branch-and-Bound and SoS theory, we devised a robust and optimal method for inlier set maximization of either point-to-plane or plane-to-plane correspondences. Although the problem is nonlinear and combinatorial, our method has provided outstanding results in terms of robustness and optimality. In particular, the employed optimization framework has the

potential to be efficiently applied to other nonlinear Computer Vision problems.

Acknowledgements This research has been funded by the International Project NRF-ANR DrAACaR: ANR-11-ISO3-0003, the Regional Council of Bourgogne and European Regional Development Fund.

References

- Bartoli, A., & Castellani, U. (2012). 3D shape registration. In *3D imaging, analysis, and applications*, Springer (pp. 221–264).
- Bazin, J., Li, H., Kweon, I. S., Démonceaux, C., Vasseur, P., & Ikeuchi, K. (2013). A branch-and-bound approach to correspondence and grouping problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 35, 1565–1576.
- Bormann, D., Elseberg, J., Lingemann, K., & Nüchter, A. (2011). The 3D hough transform for plane detection in point clouds: A review and a new accumulator design. *3D. Research*, 32(1–32), 13.
- Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. New York, NY: Cambridge University Press.
- Breuel, T. M. (2003). Implementation techniques for geometric branch-and-bound matching methods. *Computer Vision and Image Understanding*, 90(3), 258–294.
- Chandraker, M., Agarwal, S., Kahl, F., Nister, D., & Kriegman, D. (2007). Autocalibration via rank-constrained estimation of the absolute quadric. In *IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 1–8).
- Chesi, G., Garulli, A., Vicino, A., & Cipolla, R. (2002). Estimating the fundamental matrix via constrained least-squares: A convex approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 24, 397–401.
- Choi, M. D., Lam, T. Y., & Reznick, B. (1995). Sums of squares of real polynomials. *Proceedings of Symposia in Pure Mathematics*, 2(58), 103–126.
- Christy, S., & Horaud, R. (1999). Iterative pose computation from line correspondences. *Computer Vision and Image Understanding (CVIU)*, 73, 137–144.
- Corsini, M., Dellepiane, M., Ganovelli, F., Gherardi, R., Fusiello, A., & Scopigno, R. (2013). Fully automatic registration of image sets on approximate geometry. *International Journal of Computer Vision (IJCV)*, 102, 91–111.
- Du, S., Zheng, N., Ying, S., You, Q., & Wu, Y. (2007). An extension of the ICP algorithm considering scale factor. In *IEEE international conference on image processing (ICIP)* (pp. V–193).
- Enqvist, O., Josephson, K., & Kahl, F. (2009). Optimal correspondences from pairwise constraints. In *2009 IEEE 12th international conference on computer vision* (pp. 1295–1302). IEEE.
- Enqvist, O., & Kahl, F. (2008). Robust optimal pose estimation. *Computer Vision-ECCV, 2008*, 141–153.
- Ferraz, L., Binefa, X., & Moreno-Noguer, F. (2014). Very fast solution to the pnp problem with algebraic outlier rejection. In *IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 501–508).
- Finsler, P. (1936/37). Über das vorkommen definiter und semidefiniter formen in scharen quadratischer formen. *Commentarii Mathematici Helvetici*, 9, 188–192.
- Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communication of the ACM*, 24, 381–395.
- Fitzgibbon, A. W. (2001). Robust registration of 2D and 3D point sets. In *The British machine vision conference (BMVC)* (pp. 662–670).
- Fraundorfer, F., & Scaramuzza, D. (2012). Visual odometry: Part II: Matching, robustness, optimization, and applications. *IEEE Robotics & Automation Magazine*, 19(2), 78–90.
- Habed, A., Al Ismaeil, K., & Fofi, D. (2012). A new set of quartic trivariate polynomial equations for stratified camera self-calibration under zero-skew and constant parameters assumptions. In *European conference on computer vision (ECCV)* (pp. 710–723).
- Hartley, R. I., & Zisserman, A. (2004). *Multiple view geometry in computer vision* (2nd ed.). Cambridge: Cambridge University Press.
- Hilbert, D. (1888). Über die darstellung definiter formen als summe von formen quadraten. *Mathematische Annalen*, 32, 342–350.
- Hough Transform Plane Detector. (2015). Howpublished. <https://github.com/daviddoria/vtkhoughplanes/>
- Jensen, R., Dahl, A., Vogiatzis, G., Tola, E., & Aanæs, H. (2014). Large scale multi-view stereopsis evaluation. In *IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 406–413).
- Jurie, F. (1999). Solution of the simultaneous pose and correspondence problem using gaussian error model. *Computer Vision and Image Understanding*, 73(3), 357–373.
- Kahl, F., & Henron, D. (2007). Globally optimal estimates for geometric reconstruction problems. *International Journal of Computer Vision*, 74(1), 3–15.
- Lasserre, J. B. (2000). Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11, 796–817.
- Li, H. (2009). Consensus set maximization with guaranteed global optimality for robust geometry estimation. In *IEEE international conference on computer vision (ICCV)* (pp. 1074–1080).
- Liu, L., & Stamos, I. (2005). Automatic 3D to 2D registration for the photorealistic rendering of urban scenes. In *IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 137–143).
- Mastin, A., Kepner, J., & Fisher III, J. W. (2009). Automatic registration of LIDAR and optical images of urban scenes. In *IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 2639–2646).
- Moulon, P., Monasse, P., & Marlet, R. (2013). Adaptive structure from motion with a contrario model estimation. In *Asian conference on computer vision (ACCV)* (pp. 257–270).
- Olsson, C., Kahl, F., & Oskarsson, M. (2006). The registration problem revisited: Optimal solutions from points, lines and planes. In *IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 1206–1213).
- Parrilo, P. A. (2000). Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization. Technical report, California Institute of Technology
- Paudel, D. P., Démonceaux, C., Habed, A., & Vasseur, P. (2014). Localization of 2D cameras in a known environment using direct 2D–3D registration. In *International conference on pattern recognition (ICPR)* (pp. 1–6).
- Plotz, T., & Roth, S. (2015). Registering images to untextured geometry using average shading gradients. In *IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 2030–2038).
- Powers, V., & Wörmann, T. (1998). An algorithm for sums of squares of real polynomials. *Journal of Pure and Applied Algebra*, 127(1), 99–104.
- Putinar, M. (1993). Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42, 969–984.
- Ramalingam, S., & Taguchi, Y. (2013). A theory of minimal 3D point to 3D plane registration and its generalization. *International Journal of Computer Vision*, 102(1–3), 73–90.
- Rusinkiewicz, S., & Levoy, M. (2001). Efficient variants of the ICP algorithm. In *3-D digital imaging and modeling (3DIM)* (pp. 145–152).
- Schindler, G., Krishnamurthy, P., Lubliner, R., Liu, Y., & Dellaert, F. (2008). Detecting and matching repeated patterns for automatic geo-tagging in urban environments. In *IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 1–7).

- Schweighofer, G., & Pinz, A. (2008). Globally optimal $O(n)$ solution to the pnp problem for general camera models. In *The British machine vision conference (BMVC)* (pp. 1–10).
- Segal, A. V., Haehnel, D., & Thrun, S. (2009). Generalized-ICP. In *Robotics: Science and systems (RSS)*.
- Strecha, C., von Hansen, W., Van Gool, L., Fua, P., & Thoennessen, U. (2008). On benchmarking camera calibration and multi-view stereo for high resolution imagery. In *IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 1–8).
- Tamaazousti, M., Gay-Bellile, V., Collette, S. N., Bourgeois, S., & Dhome, M. (2011). Nonlinear refinement of structure from motion reconstruction by taking advantage of a partial knowledge of the environment. In *IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 3073–3080).
- Taneja, A., Ballan, L., & Pollefeys, M. (2013). City-scale change detection in cadastral 3D models using images. In *IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 113–120).
- Verschelde, J. (1999). Algorithm 795: Phcpack: A general-purpose solver for polynomial systems by homotopy continuation. *ACM Transactions on Mathematical Software (TOMS)*, 25(2), 251–276.
- Viola, P., & Wells, M. I. I. (1997). Alignment by maximization of mutual information. *International Journal of Computer Vision (IJCV)*, 24, 137–154.
- Wagner, S. (2009). Archimedean quadratic modules: A decision problem for real multivariate polynomials. Ph.D. thesis, Universität Konstanz.
- Yang, J., Li, H., & Jia, Y. (2013). Go-icp: Solving 3D registration efficiently and globally optimally. In *IEEE international conference on computer vision (ICCV)* (pp. 1457–1464).
- Yang, J., Li, H., & Jia, Y. (2014). Optimal essential matrix estimation via inlier-set maximization. In *European conference on computer vision (ECCV)* (pp. 111–126).
- Zhang, X., Agam, G., & Chen, X. (2014). Alignment of 3d building models with satellite images using extended chamfer matching. In *IEEE conference on computer vision and pattern recognition workshops (CVPRW)* (pp. 746–753).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.