Università degli studi di Udine

Robust synchronization in SO(3) and SE(3) via low-rank and sparse matrix decomposition

Original

Availability:
This version is available http://hdl.handle.net/11390/1145816 since 2020-03-04T08:57:23Z

Publisher:

Published
DOI:10.1016/j.cviu.2018.08.001

Terms of use:
The institutional repository of the University of Udine (http://air.uniud.it) is provided by ARIC services. The aim is to enable open access to all the world.

Publisher copyright

(Article begins on next page)

12 October 2024

Computer Vision and Image Understanding
journal homepage: www.elsevier.com

Robust synchronization in SO(3) and SE(3) via low-rank and sparse matrix decomposition

Federica Arrigoni[a], Beatrice Rossi[b], Pasqualina Fragneto[b], Andrea Fusiello[a,**]

[a]*DPIA - University of Udine, Via delle Scienze 208, Udine 33100, Italy*
[b]*STMicroelectronics, Via Olivetti 2, Agrate Brianza 20864, Italy*

ABSTRACT

This paper deals with the synchronization problem, which arises in multiple 3D point-set registration and in structure-from-motion. The problem is formulated as a *low-rank and sparse* matrix decomposition that caters for missing data, outliers and noise, and it benefits from a wealth of available decomposition algorithms that can be plugged-in. A minimization strategy, dubbed R-GoDec, is also proposed. Experimental results on simulated and real data show that this approach offers a good trade-off between resistance to outliers and speed.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

This paper deals with the synchronization (a.k.a. averaging) problem that arises in Computer Vision in the context of structure-from-motion and multiple 3D point-set registration. The goal is to compute a set of direct isometries that represent the *absolute* position and angular attitudes of a set of local reference frames, given their *relative* positions and angular attitudes. These local frames can be camera reference frames, in which case we are in the context of structure-from-motion, or local coordinates where 3D points are represented, in which case we are dealing with a 3D point-set registration problem. A related problem restricts the attention to the angular attitude, leaving out the position, which is analysed in depth by Hartley et al. (2013) under the name "multiple rotation averaging".

More abstractly, the aim of *group synchronization* (Giridhar and Kumar, 2006; Singer, 2011) is to recover elements of a group from noisy measures of their ratios. In our case, when considering the full rigid transformations linking the local frames, we set ourself to work in the group of direct isometries, or Special Euclidean Group SE(3), which is the semi-direct product of the Special Orthogonal Group SO(3) with $\mathbb{R}^3$. Likewise, when only the angular attitudes are considered, they are represented by rotations, *i.e.*, elements of the Special Orthogonal Group SO(3).

Our solution to these two synchronization problems is inspired by recent advances in the fields of *robust principal component analysis* (RPCA) and *matrix completion* (MC). The main and original contribution of this paper is the formulation of synchronization as a "low-rank and sparse" (LRS) matrix decomposition, by designing a novel cost function that naturally includes missing data and outliers in its definition. In principle, any LRS decomposition method able to deal with outliers and missing data can be plugged into this framework, such as Grasta (He et al., 2012) or L1-Alm (Zheng et al., 2012).

In this respect, our approach could benefit from any development in the field of LRS matrix decomposition, such as the R-GoDec algorithm that we conceive here as an extension of the GoDec algorithm (Zhou and Tao, 2011) that is able to address RPCA and MC simultaneously.

Exhaustive experiments on synthetic and real data are performed to evaluate the proposed approach against state-of-the-art algorithms. Results show that our formulation can be profitably applied to structure-from-motion and multiple 3D point-set registration, it is computationally undemanding and provides a good trade-off between statistical efficiency and resilience to outliers. A drawback of the LRS approach is that it is more affected than others by the sparsity of the measurement matrix. It is worth to remember, though, that the goal of synchronization is to exploit redundancy: if the measures are barely sufficient the problem starts loosing significance.

The paper is organized as follows. Applications of the synchronization problem are presented in Section 2 while existing

**Corresponding author.
e-mail: andrea.fusiello@uniud.it (Andrea Fusiello)

solutions are described in Section 3. Section 4 is a comprehensive introduction to LRS decomposition, which can be regarded as one of the contributions of this paper. Section 5 defines the synchronization problem, whereas Section 6 describes how such a problem can be translated into a LRS decomposition of an incomplete matrix, corrupted by noise and outliers. The approach proposed in this paper is supported by experimental results on both synthetic and real data, reported in Sections 7 and 8. The conclusions are drawn in Section 9.

This work is an extended version of (Arrigoni et al., 2014a), where the LRS approach was introduced and applied to SO(3), and (Arrigoni et al., 2016b), where it has been applied to SE(3). In this paper the R-GoDec algorithm is extended from ℓ_1 to mixed $\ell_{2,1}$-norm, in order to reflect the block structure of the measures, and a discussion on its convergence is included. It also contains a more thorough description of the background, and further experiments that cover additional algorithms and datasets.

2. Applications

In this section we shall describe in some details how synchronization in SO(3) and SE(3) can be applied to structure-from-motion and multiple 3D point-set registration. Other applications briefly surveyed here include sensor network localization, cryo-electron microscopy, and simultaneous localization and mapping.

2.1. Multiple 3D point-set registration

The goal of multiple 3D point-set registration is to find the rigid transformations that bring multiple ($n \geqslant 2$) 3D point sets into alignment, where each rigid transformation is represented by a direct isometry, *i.e*. one element of SE(3). Such point sets usually come from a 3D scanning device, which can frame only a fraction of an object from a given viewpoint. Therefore, registration of multiple scans is necessary to build a full 3D model of the object. This problem covers a wide range of applications, including (but not limited to) cultural heritage, reverse engineering and virtual reality.

If $n = 2$ then we are dealing with a *pairwise* (two point-sets) registration problem. The gold standard in this context is the Iterative Closest Point (ICP) Algorithm (Besl and McKay, 1992; Chen and Medioni, 1991), which computes correspondences between the point sets given an estimate for the rigid transformation, then updates the transformation based on the current correspondences, and iterates through these steps until convergence – to a local minimum – is reached. See (Rusinkiewicz and Levoy, 2001) for several variants of the ICP Algorithm.

If $n > 2$ then we are dealing with a *multiple* point-set registration problem, which is more complex than the $n = 2$ case due to the high amount of parameters that have to be estimated. Among the early efforts to address this problem are the *sequential* approaches proposed by Chen and Medioni (1991) and Pulli (1999), that repeatedly register two point sets and integrate them into one model, until all the scans are considered. This approach however does not take into account all the available constraints, *e.g*. the constraint between the last and first point set is not used if the sets are obtained using a turntable. Therefore the solution is suboptimal.

A different paradigm is realized by *global* methods, which are able to register simultaneously all the points sets. Such techniques take advantage of the redundancy in relative motions by using *all* the constraints available between pairs of scans, thus they distribute the errors evenly across the scans, preventing drift in the solution.

Global registration can be solved in point (correspondences) space or in frame space. In the first case, all the rigid transformations are simultaneously optimized with respect to a cost function that includes the distance between corresponding points (Pennec, 1996; Benjemaa and Schmitt, 1998; Krishnan et al., 2007; Toldo et al., 2010; Bonarrigo and Signoroni, 2011; Fantoni et al., 2012; Chaudhury et al., 2015).

In the second case the optimization criterion is related to the internal coherence of the network of rotations (and translations) applied to the local coordinate frames. Early frame-space methods include (Fusiello et al., 2002) and (Sharp et al., 2002). More recent works comprise (Torsello et al., 2011; Govindu and Pooja, 2014; Bernard et al., 2015; Arrigoni et al., 2016c). Frame-space methods are faster and less memory-demanding than point-space ones. Unquestionably, any optimal formulation *must* include points in the cost function, in analogy to bundle adjustment in the context of structure-from-motion. Nevertheless, frame-space approaches yield a fairly accurate registration.

The formulation by Shih et al. (2008) lies at the boundary between frame-space and point-space methods, where 3D points are used to compute a second-order approximation of the cost function, but they are not involved in subsequent computations.

2.2. Structure-from-motion

Recovering geometric information about a scene captured by multiple cameras has a great relevance in Computer Vision. In the structure-from-motion problem such geometric information includes both scene structure, *i.e*., 3D coordinates of scene points, and camera motion, *i.e*., absolute positions and attitudes of the cameras. This problem also appears in the context of Photogrammetry under the name of *block orientation*. Several systems have been developed to reconstruct large-scale scenes from a collection of unordered images, recently surveyed by Ozyesil et al. (2017). They can be divided into three categories: *structure-first*, *structure-and-motion*, *motion-first*.

Structure-first approaches (*e.g*., Crosilla and Beinat (2002)) begin with estimating the structure and then compute the motion. Specifically, stereo-models are built and co-registered, similarly to the multiple 3D point-set registration problem.

Structure-*and*-motion techniques solve simultaneously for structure and motion. Bundle block adjustment (Triggs et al., 2000; Fusiello and Crosilla, 2015), resection-intersection methods (Snavely et al., 2006; Brown and Lowe, 2005), and hierarchical methods (Gherardi et al., 2010; Ni and Dellaert, 2012) belong to this category. Although being highly accurate, these approaches suffer from two main disadvantages: on one hand they require intermediate expensive non-linear minimizations to damp error propagation, on the other hand the output may

depend on the order in which images are added or on the choice of the initial pair/triplet.

Motion-first methods initially recover the motion and then compute the structure (Govindu, 2001; Martinec and Pajdla, 2007; Arie-Nachimson et al., 2012; Moulon et al., 2013; Jiang et al., 2013; Ozyesil and Singer, 2015; Goldstein et al., 2016; Arrigoni et al., 2016a). They start from the relative motions determined from point matches among the images, they compute the angular attitude and position of the cameras with respect to an absolute coordinate frame, and then they return a sparse 3D point cloud representing the scene. These motion-first methods are *global*, for they take into account the entire relative information at once, or, in other terms, they consider the whole *epipolar graph* (also known as the *viewing graph*), where the nodes correspond to the cameras and the edges represent epipolar relationships. Global techniques have the advantage of fairly distributing the errors among the cameras, and thus they need bundle adjustment only at the end, thereby resulting in a reduction of the computational cost. For this reason they have gained increasing attention in the community, similarly to frame-based methods for multiple 3D point-set registration.

Since the magnitude of the relative translations is not available, structure-from-motion cannot be solved straightforwardly as a synchronization in SE(3). Most techniques split the motion estimation process in two stages: a rotation synchronization to obtain the angular attitudes, followed by the recovery of camera positions, which in turn can be cast as a synchronization in $\mathbb{R}^3$ after computing the magnitude of relative translations (Arrigoni et al., 2015), or as a bearing-only network localization (Zhao and Zelazo, 2016) using only the directions of the relative displacements. The direct computation of translation magnitudes allows also to address the problem in SE(3), as in (Govindu, 2004; Arrigoni et al., 2016c).

As far as the structure-from-motion is concerned, we will concentrate here only on the rotation synchronization stage. This application has been recently surveyed by Tron et al. (2016).

2.3. Other applications

The synchronization problem also arises in the context of sensor network localization. In such a scenario the nodes of a sensor network can measure each other's relative rotation (by means of *e.g.* angle-of-arrival) with respect to their relative reference frames, and the goal is to express some other sensor measurements in a unique/global reference frame (measurements might include positions of targets, environment elements, etc.). Usually this application refers to planar networks, namely the synchronization problem in SO(2) (Cucuringu et al., 2012; Piovan et al., 2013).

Another application regards structural biology. In (Singer and Shkolnisky, 2011) the problem of recovering the three-dimensional structure of a macromolecule from many cryo-electron microscopy (cryo-EM) images is considered. The direction from which each image is taken is unknown, and a rotation synchronization technique is used for determining the viewing direction of all cryo-EM images at once.

The synchronization problem is also related to simultaneous localization and mapping (SLAM), where the goal is to orient a robot moving in an unknown environment, while building a map of the environment (Carlone et al., 2015; Rosen et al., 2015; Bourmaud, 2016). This task can be reduced to a synchronization in SE(3), if measures of relative poses are available.

3. Related work

Several approaches have been proposed to solve the synchronization problem in SO(3) and SE(3), in the context of the applications surveyed in the previous section. We shall divide them into non-robust and robust methods, according to the resilience they show to rogue measures. In general, robustness is gained at the expense of statistical efficiency, *i.e.*, non-robust estimators typically get closer to the Cramèr-Rao bound (Boumal et al., 2014). On the other hand, non-robust methods can be skewed even by a single outlier, hence they are not applicable in practice unless they are preceded by an outlier detection stage.

Outliers are ubiquitous in the real world. In the structure-from-motion context, for example, repetitive structures in the images cause mismatches which wreck the epipolar geometry. In multiple 3D point-set registration, outliers are caused by faulty pairwise registration, which in turn may be originated by insufficient overlap or poor initialization.

3.1. Non-robust methods

Sharp et al. (2002) decompose the viewing graph into a set of cycles, and they propose an iterative procedure to recover the absolute rotations (and translations) in which the error is distributed over these cycles. The same idea appeared also in (Dubbelman and Browning, 2015; Peters et al., 2015). As observed by Govindu and Pooja (2014), this technique performs a suboptimal set of averages and as a result it may converge to a local minimum.

In (Govindu, 2001; Fusiello et al., 2002; Fredriksson and Olsson, 2012) the synchronization problem is cast to the optimization of an objective function where rotations are parametrized as unit quaternions. Govindu (2001) expresses the compatibility constraint between relative and absolute rotations as a linear system of equations which is solved in the least-squares sense, while in (Fusiello et al., 2002) the absolute rotations (and translations) are computed using a quasi-Newton method. Fredriksson and Olsson (2012) exploit the theory of Lagrangian duality and develop a procedure to verify the global optimality of a local solution to rotation synchronization. Similar to the way that rotations can be represented by quaternions, rigid motions can be represented by dual quaternions. This parametrization is used in (Torsello et al., 2011) where a diffusion algorithm is proposed that updates each absolute transformation in turn through linear or geodesic averaging.

The methods described in (Martinec and Pajdla, 2007; Arie-Nachimson et al., 2012; Arrigoni et al., 2014b) perform ℓ_2 averaging of relative rotations by using the chordal (Frobenius) distance. Without enforcing the orthogonality constraints, approximate solutions are computed, and they are subsequently projected onto SO(3) by finding the nearest rotation matrices (in the Frobenius norm sense). Martinec and Pajdla (2007) compute a least-squares solution through vectorization and Singular

Value Decomposition (SVD), and this approach is extended by Arie-Nachimson et al. (2012) using spectral decomposition or semi-definite programming. A gradient descent method based on matrix completion is presented in (Arrigoni et al., 2014b). According to the analysis in (Martinec and Pajdla, 2007), methods involving matrices usually perform better than quaternion minimization.

The methods presented in (Bernard et al., 2015; Arrigoni et al., 2016c; Rosen et al., 2015, 2016) perform synchronization in SE(3) using the matrix representation of the Special Euclidean group combined with an ℓ_2 cost function. After relaxing the geometric constraints of rigid motions, approximate solutions are derived, which are then projected onto SE(3) (in the Frobenius norm sense). In (Bernard et al., 2015; Arrigoni et al., 2016c) a closed-form solution is found via eigenvalue decomposition, which can be regarded as the extension to SE(3) of the spectral synchronization in SO(3) (Arie-Nachimson et al., 2012). In (Rosen et al., 2015) the Special Euclidean group is relaxed to its convex hull, which has a semi-definite representation (Saunderson et al., 2015). A convex relaxation is also employed in (Rosen et al., 2016) where the authors, using the theory of Lagrangian duality, develop an algorithm for certifying the global optimality of a candidate solution to synchronization in SE(3).

In (Tron and Danilidis, 2014) the Riemannian manifold structure of SE(3) is exploited and the absolute transformations are found via Riemannian gradient descent. Non isotropic noise and incomplete measurements are taken into account through the use of covariance matrices. However, this is a local iterative method, hence it requires a good initialization.

A different approach is followed by Govindu (2004), where the Lie-group structure of SO(3) and SE(3) is exploited, and an iterative scheme is proposed in which at each step the absolute rotations (and translations) are updated by averaging two-view transformations in the tangent space. Originally proposed in the structure-from-motion framework, the same technique was also applied to simultaneous localization and mapping (Agrawal, 2006) and multiple 3D point-set registration (Govindu and Pooja, 2014).

3.2. Robust methods

The main drawback of the previous techniques is that they suffer from the presence of inconsistent relative rotations/translations, and thus they need a preliminary step to detect and remove such outliers *before* computing the absolute rotations/translations.

A comprehensive survey of methods aimed at detecting outliers can be found in (Moulon et al., 2013). These approaches check for cycle consistency, *i.e.*, deviation from identity, within the epipolar graph (Enqvist et al., 2011; Arrigoni et al., 2014b; Zach et al., 2010; Bourmaud et al., 2014; Govindu, 2006; Olsson and Enqvist, 2011). Enqvist et al. (2011) consider a maximum-weight spanning tree, where the weight of an edge is the number of inlier correspondences, and they analyse cycles formed by the remaining edges. In (Arrigoni et al., 2014b) some heuristics based on cycle bases are introduced to improve this scheme. In (Zach et al., 2010) a Bayesian framework is used to classify all the edges of the epipolar graph into inliers and outliers. Moulon et al. (2013) showed that an iterative use of this method, adjusted with the cycle length weighting by Enqvist et al. (2011), can remove most outlier edges in the graph. In (Bourmaud et al., 2014) it is assumed that there exists a spanning tree without dependent outliers (it may contain independent outliers), and an iterative approach based on a Kalman filter is developed for outlier detection. Other approaches (Govindu, 2006; Olsson and Enqvist, 2011) are based on random spanning trees, in a RANSAC-like fashion.

These strategies are computationally demanding and do not scale well with the number of cameras. For example, Olsson and Enqvist (2011) reports that, after feature extraction and matching, outlier removal is the most expensive step within the entire structure-from-motion pipeline.

Recently, some approaches have been developed to robustly solve the synchronization problem *without* detecting outliers explicitly. Techniques in (Hartley et al., 2011; Chatterjee and Govindu, 2013; Wang and Singer, 2013; Crandall et al., 2011; Tron and Vidal, 2014; Arrigoni et al., 2016c; Boumal et al., 2013), together with the approach presented in this paper, fall under this category.

Tron and Vidal (2014) compute the absolute rotations via Riemannian gradient descent. The choice of the step-size and several cost functions are discussed, including a reshaped cost function, which is similar to robust M-estimators and is less sensitive to large errors. In (Boumal et al., 2013) a statistical approach is adopted by assuming a specific noise model that takes into account also the presence of outliers, and a maximum likelihood estimator is computed via Riemannian trust-region optimization.

In (Hartley et al., 2011; Wang and Singer, 2013) a cost function based on the ℓ_1 norm is used to average relative rotations, exploiting the fact that the ℓ_1 norm is more robust to outliers than the ℓ_2 norm. In (Hartley et al., 2011) the geodesic (angular) distance is used, while in (Wang and Singer, 2013) the chordal metric is adopted. Wang and Singer (2013) consider a semidefinite relaxation and use the alternating direction augmented Lagrangian method to minimize the cost function. They focus on accuracy rather than efficiency, providing theoretical results about exact and stable recovery of rotations. In (Hartley et al., 2011) each absolute rotation is updated in turn by applying the Weiszfeld algorithm to its neighbours. This technique is generalized to ℓ_q optimization in (Aftab et al., 2015), with $1 \leqslant q < 2$, where improved reliability and robustness is shown compared to using the ℓ_2-norm.

Crandall et al. (2011) use a truncated quadratic as a more robust self-consistency error. This method uses a discrete Markov random field formulation, combined with a continuous Levenberg-Marquardt refinement. In addition to relative motions, vanishing points and information from other sensors are assumed as input.

As observed by Chatterjee and Govindu (2013), neither (Crandall et al., 2011) nor the Weiszfeld algorithm satisfies both the requirements of a computationally efficient and scalable robust scheme. On one hand, the Weiszfeld method scales poorly with large datasets, since any change in a given rotation takes

a long time to propagate over the entire graph. On the other hand, the method by Crandall et al. (2011) can handle large-scale problems, but it requires a significant amount of memory.

To overcome these drawbacks, Chatterjee and Govindu (2013) proposed a two-stage synchronization scheme that extends the Lie-averaging algorithm by Govindu (2004). First, the ℓ_1-norm of a vector that contains both noise and outliers is minimized, exploiting recent work in compressed sensing. Then, this solution is improved through iteratively reweighted least squares (IRLS). Chatterjee and Govindu (2017) proved that the method in (Chatterjee and Govindu, 2013) is a quasi-Newton optimization and analysed different robust loss functions in the IRLS step. Experiments in (Chatterjee and Govindu, 2013, 2017) demonstrate that such technique is a robust solution to rotation synchronization even for large-scale datasets.

An IRLS approach is also adopted in (Arrigoni et al., 2016c) for synchronization in SE(3), based on the observation that the spectral solution can be easily extended to handle weighted measurements, where the weights reflect the reliability of the pairwise measurements. Experimental results reported in (Arrigoni et al., 2016c) show that this method produces high-quality results in low execution time.

4. Low-rank and sparse matrix decomposition

Matrix decompositions have a long history and often occur in the analysis of complex data. The idea is that decomposing a data matrix into the sum of terms with specific properties makes the understanding easier as it separates information into simpler pieces. In recent years, decompositions imposing constraints on the rank and sparsity of the addends have become very popular thanks to their profitable application in several fields, such as pattern recognition, machine learning, and signal processing.

Let $\widehat{X}$ be a data matrix, and suppose that $\widehat{X}$ is known to be the exact or approximate sum of a low-rank term and a sparse term. Low-rank and sparse (LRS) decompositions address problems of the general form

$$\mathcal{F}(\widehat{X}) = \mathcal{F}(L) + S + N \tag{1}$$

where $\mathcal{F}$ is a linear operator, L is an unknown low-rank matrix, S is an unknown sparse matrix and N is a diffuse noise. Generally, the sparse term S represents gross errors affecting the measurements (outliers), while the low-rank part represents some meaningful low-dimensional structure contained into the data. The goal is to recover L (and possibly S) under different conditions for S, N and $\mathcal{F}$. A survey on this topic is reported in (Zhou et al., 2014).

4.1. Robust Principal Component Analysis

An example of LRS decomposition is Robust Principal Component Analysis (RPCA) (Candès et al., 2011). The goal is to find the lowest-rank matrix L and the sparsest matrix S such that a given data matrix $\widehat{X}$ can be decomposed as

$$\widehat{X} = L + S + N \tag{2}$$

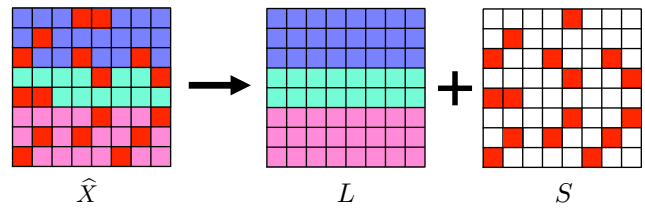


Figure 1: Robust Principal Component Analysis. S is the outlier term.

with N a diffuse noise. This is illustrated in Figure 1. Please observe that such a decomposition is an instance of the general problem (1) with $\mathcal{F}$ being the identity operator.

A suitable minimization problem for RPCA is

$$\begin{cases} \min\limits_{L,S} \|L\|_* + \lambda \|S\|_1 \\ \text{s.t.} \ \left\|\widehat{X} - L - S\right\|_F \leqslant \epsilon \end{cases} \tag{3}$$

where $\|\cdot\|_*$ denotes the nuclear norm, $\|\cdot\|_F$ denotes the Frobenius norm, $\|S\|_1$ is the ℓ_1-norm of S (viewed as a vector), and ϵ, λ are given parameters. It is well known from sparse representation theory that minimizing the ℓ_1-norm promotes sparse vectors (Fornasier, 2010). Moreover, the nuclear norm is the tightest convex relaxation of the rank function (Fazel, 2002), since it is the *sum* of the singular values of a matrix. Thus the solution of problem (3) is expected to recover a blind separation between the lowest-rank component and the sparsest errors contained into the data, *i.e.*, the outliers.

Theoretical conditions under which such a solution is stable with respect to a diffuse noise N with high probability are studied in (Zhou et al., 2010) and they depend on some incoherence properties of the data matrix and on the sparsity pattern of S.

Available algorithms for RPCA include, among others, the Accelerated Proximal Gradient (APG) method (Zhou et al., 2010) and extensions of the Augmented Lagrange Multiplier (ALM) method such as (Lin et al., 2010) or the ASALM algorithm (Tao and Yuan, 2011). These approaches however involve repeated computation of the SVD (or at least of a partial SVD) of matrices of considerable size which represents the principal bottleneck of current solutions for RPCA.

A faster alternative to RPCA is the randomized approximate matrix decomposition (Tropp et al., 2010). This approach is based on the observation that the low-rank term L of a decomposition of the form (2) can be well approximated by random projections onto its column space, thus providing a fast approximation of SVD.

A technique exploiting this paradigm is the GoDec algorithm described in (Zhou and Tao, 2011). This method requires to know approximately both the rank r of the low-rank term L and the cardinality (*i.e.*, the number of non-zero entries) k of the sparse term S, and it solves the following minimization problem

$$\begin{cases} \min\limits_{L,S} \left\|\widehat{X} - L - S\right\|_F^2 \\ \text{s.t. } \operatorname{rank}(L) \leqslant r, \ \ \operatorname{card}(S) \leqslant k. \end{cases} \tag{4}$$

GoDec adopts a *block-coordinate* minimization scheme

(a.k.a. *block relaxation* or *alternating optimization*), *i.e.*, it alternatively forces L to the rank-r approximation of $\widehat{X} - S$, and forces S to the sparse approximation with cardinality k of $\widehat{X} - L$. The rank-r projection is computed using Bilateral Random Projections (BRP) instead of SVD thus obtaining a speed up in the computation. The updating of S is obtained via entry-wise hard thresholding, keeping the k largest elements of $|\widehat{X} - L|$ only.

Estimating the cardinality k of the sparse term might be unreliable in practical applications. In order to avoid this parameter, one can consider instead the following minimization problem:

$$\begin{cases} \min\limits_{L,S} \dfrac{1}{2} \left\| \widehat{X} - L - S \right\|_F^2 + \lambda \left\| S \right\|_1 \\ \text{s.t. } \operatorname{rank}(L) \leqslant r \end{cases} \tag{5}$$

where λ is a regularization parameter which balances the tradeoff between the sparsity of S and the residual error $\left\| \widehat{X} - L - S \right\|_F^2$. In this case, the updating of the sparse part is obtained by minimizing the cost function in (5) with respect to S, keeping L constant. Such a problem is known to have an analytical solution, given by the *soft thresholding* or *shrinkage* operator Θ_λ (Beck and Teboulle, 2009) applied to the matrix $\widehat{X} - L$. This operator is defined as follows

$$\Theta_\lambda(S) = \operatorname{sign}(S) \cdot \max(0, |S| - \lambda) \tag{6}$$

where scalar operations are applied element-wise. This method is described in detail in Algorithm 1.

Algorithm 1 GoDec for RPCA

Input: $\widehat{X}, r, \epsilon, \lambda$
Output: L, S
Initialize: $L = \widehat{X}, S = 0$
while $\left\| \widehat{X} - L - S \right\|_F^2 / \left\| \widehat{X} \right\|_F^2 > \epsilon$ **do**
1. $L \leftarrow$ rank-r approximation of $\widehat{X} - S$ via BRP
2. $S \leftarrow \Theta_\lambda(\widehat{X} - L)$

end while

A principled choice of λ, which plays a role similar to an inlier threshold, is derived in (Donoho, 1995) in the case of uncorrelated residuals

$$\lambda = \sigma \sqrt{2 \log(m)} \tag{7}$$

where m is the number of observations and σ is an estimate of the noise standard deviation (we used the default value of 0.02 in all our experiments).

4.2. Matrix completion

RPCA assumes that the data matrix $\widehat{X}$ is fully available. However, in practical scenarios, one has to face the problem of missing data. Matrix Completion (MC) (Candès and Recht, 2009; Candès and Tao, 2010) is the most natural tool to manage matrices containing unspecified entries.

A *partial* matrix is a matrix whose entries are specified on a subset of index pairs and unspecified elsewhere; a *completion* of a partial matrix consists in assigning values to the unspecified entries. Matrix completion problems deal with partial matrices which satisfies some prescribed properties, notably low-rank or positive definiteness. We are concerned here with the low-rank problem, illustrated in Figure 2, which can be cast as an instance of the general decomposition (1) with a specific choice of $\mathcal{F}$ and $S = 0$, namely

$$\mathcal{P}_\Omega(\widehat{X}) = \mathcal{P}_\Omega(L) + N. \tag{8}$$

Here Ω is a $(0,1)$-matrix representing the *pattern* (a.k.a. *sampling set*) of $\widehat{X}$, *i.e.*, $\Omega_{ij} = 1$ if $\widehat{X}_{ij}$ is specified and $\Omega_{ij} = 0$ otherwise, and $\mathcal{P}_\Omega(X) = \Omega \circ X$, where $\circ$ is the Hadamard (entry-wise) product, with the provision that an unspecified value multiplied by 0 gives 0.

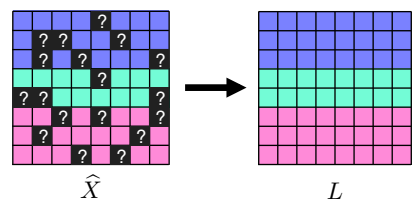


Figure 2: L is a low rank completion of $\widehat{X}$.

The MC problem can be solved through nuclear norm minimization

$$\begin{cases} \min\limits_{L} \left\| L \right\|_* \\ \text{s.t. } \left\| \mathcal{P}_\Omega(\widehat{X} - L) \right\|_F \leqslant \epsilon \end{cases} \tag{9}$$

or – if the rank is known a priori – by addressing the following optimization problem

$$\begin{cases} \min\limits_{L} \left\| \mathcal{P}_\Omega(\widehat{X} - L) \right\|_F^2 \\ \text{s.t. } \operatorname{rank}(L) \leqslant r. \end{cases} \tag{10}$$

The minimum sample size required to make matrix completion well-posed is equal to the number of degrees of freedom of the data matrix, namely $(n_1 + n_2 - r)r$, where r is the rank and $n_1 \times n_2$ is the dimension of the matrix. In general, practical MC algorithms do not cope with the minimal case but they require a redundant number of observed entries. For instance, nuclear norm minimization recovers the full low-rank matrix L with high probability – under suitable assumptions – if the number of observed entries is of the order $O(\bar{n} r \log^6(\bar{n}))$, where $\bar{n} = \max\{n_1, n_2\}$ (Candès and Tao, 2010). Similarly to RPCA, such theoretical conditions depend on some incoherence properties of the data matrix and on the randomness of Ω. With reference to Problem (10), Keshavan et al. (2010) improved the bound on the cardinality of Ω to $O(\bar{n} r \log(\bar{n}))$, with the extra condition that the data matrix has bounded condition number.

Conventional solvers for MC include convex solvers such as ALM (Lin et al., 2010), SVT (Cai et al., 2010) and FPCA (Ma et al., 2011), and subspace identification solvers such as OptSpace (Keshavan et al., 2010) and ADMiRA (Lee and Bresler, 2010).

Specifically, in the subspace identification problem the goal is to identify the column space of the unknown low-rank term L. Clearly, any matrix L of rank up to r admits a factorization of the form $L = ZY^T$ where Z and Y^T are of r columns and r rows respectively. Thus an alternative minimization for the MC problem is

$$\begin{cases} \min\limits_{L,Z,Y} \|ZY^T - L\|_F^2 \\ \text{s.t. } \mathcal{P}_\Omega(\widehat{X}) = \mathcal{P}_\Omega(L). \end{cases} \tag{11}$$

In particular, OptSpace solves a normalized version of the previous problem, with Z, Y belonging to the Grassmannian manifold, namely the set of all r-dimensional subspaces of a Euclidean space, via gradient descent.

The MC problem can also be solved by modifying the GoDec Algorithm, as explained by Zhou and Tao (2011). The minimization problem (10) is reformulated by introducing a sparse term S which approximates $-\mathcal{P}_\mho(L)$, where $\mho$ represents the complementary of Ω, resulting in the following problem

$$\begin{cases} \min\limits_{L,S} \left\|\mathcal{P}_\Omega(\widehat{X}) - L - S\right\|_F^2 \\ \text{s.t. } \operatorname{rank}(L) \leqslant r,\ \operatorname{supp}(S) = \mho \end{cases} \tag{12}$$

where $\operatorname{supp}(S)$ denotes the support of S, *i.e.*, the $(0,1)$-matrix with ij-th entry equal to 1 if $S_{ij} \neq 0$, and equal to 0 otherwise. The associated decomposition problem is

$$\mathcal{P}_\Omega(\widehat{X}) = L + S + N \tag{13}$$

which is equivalent to (8) but it does not involve the projection operator $\mathcal{P}_\Omega$ in the right side, thanks to the introduction of the auxiliary variable S. Note that here S does not represent the outliers, but the recovery of missing entries. In the GoDec algorithm for MC, the updating of the sparse term is obtained by assigning $\mathcal{P}_\mho(\widehat{X} - L) = -\mathcal{P}_\mho(L)$ to S. The method is summarized in Algorithm 2.

Algorithm 2 GoDec for MC

Input: $\widehat{X}, \Omega, r, \epsilon$
Output: L, S
Initialize: $L = \widehat{X}, S = 0$
while $\left\|\mathcal{P}_\Omega(\widehat{X}) - L - S\right\|_F^2 / \left\|\mathcal{P}_\Omega(\widehat{X})\right\|_F^2 > \epsilon$ **do**
1. $L \leftarrow$ rank-r approximation of $\mathcal{P}_\Omega(\widehat{X}) - S$ via BRP
2. $S \leftarrow -\mathcal{P}_\mho(L)$

end while

4.3. RPCA and MC

Although being two instances of the same general formulation (1), RPCA and MC remain two distinct problems. On one hand, RPCA handles the presence of outlier measurements but it does not deal with missing data, on the other hand MC techniques can fill missing entries, but they are not robust to outliers. Addressing these issues simultaneously is equivalent to solving the following decomposition problem (illustrated in Figure 3)

$$\mathcal{P}_\Omega(\widehat{X}) = \mathcal{P}_\Omega(L) + S + N \tag{14}$$

which aims at recovering the low-rank matrix L starting from an incomplete subset of its entries which are corrupted by both noise and outliers, where S and N have support in Ω. This problem is also known as *robust matrix completion*.

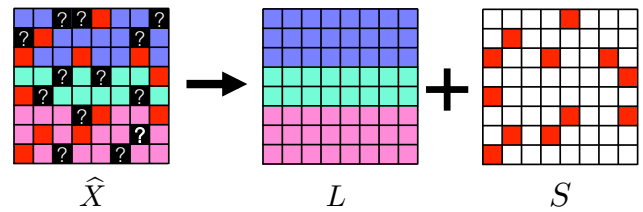


Figure 3: Robust Matrix Completion. L is a low rank completion of $\widehat{X} - S$.

This problem is numerically challenging and poorly studied from a theoretical point of view, as confirmed by the analysis in (Zhou et al., 2014). A seminal work is presented in (Waters et al., 2011), where the authors combine a greedy pursuit for updating the sparse term, with an SVD-based approximation for the low-rank term. This method requires to know in advance the cardinality of the sparse term. Other available approaches are (Tao and Yuan, 2011), which reformulates the problem under the scope of the classical ALM, and (He et al., 2012; Wen et al., 2012; Zheng et al., 2012; Wang et al., 2014) which exploit a different formulation in terms of subspace identification in the presence of outliers.

In particular, the Grasta algorithm presented by He et al. (2012) minimizes the following cost function

$$\begin{cases} \min\limits_{S,Z,Y} \|S\|_1 \\ \text{s.t. } \mathcal{P}_\Omega(\widehat{X}) = \mathcal{P}_\Omega(ZY^T) + S, \end{cases} \tag{15}$$

with Z belonging to the Grasmannian manifold. Grasta works on one column of $\widehat{X}$ at a time, *i.e.*, it considers the following minimization problem

$$\begin{cases} \min\limits_{s,Z,y} \|s\|_1 \\ \text{s.t. } \mathcal{P}_\Omega(\widehat{x}) = \mathcal{P}_\Omega(Zy^T) + s, \end{cases} \tag{16}$$

where $\widehat{x}$, s and y are column vectors of $\widehat{X}, S$ and Y respectively. The Augmented Lagrangian of this constrained minimization problem is

$$\begin{aligned} \mathcal{L}(Z,s,y,w) = \|s\|_1 + w^T(\mathcal{P}_\Omega(Zy^T) + s - \mathcal{P}_\Omega(\widehat{x})) \\ + \frac{\rho}{2}\left\|\mathcal{P}_\Omega(Zy^T) + s - \mathcal{P}_\Omega(\widehat{x})\right\|^2 \end{aligned} \tag{17}$$

where w is the dual vector. Grasta alternates between estimating Z and the triple of vectors (s,y,w). For computing Z, Grasta uses gradient descent on the Grasmannian with (s,y,w) fixed. With Z fixed, the triple (s,y,w) is computed using the Alternating Direction Method of Multipliers (ADMM) (Boyd et al., 2011).

The L1-Alm algorithm presented by Zheng et al. (2012) exploits a similar approach and solves instead

$$\begin{aligned} \min_{Z,Y} \quad \left\|\mathcal{P}_\Omega(\widehat{X} - ZY^T)\right\|_1 + \lambda \left\|Y^T\right\|_* \\ \text{s.t. } Z^T Z = I_r. \end{aligned} \tag{18}$$

Enforcing Z to be column orthogonal shrinks the solution space, while the (convex) nuclear norm regularization term is introduced to improve convergence. The optimization problem is solved via the augmented Lagrange multiplier (ALM) method (Bertsekas, 1982). At each iteration, the augmented Lagrange function with orthogonal Z is minimized using a Gauss-Seidel strategy, then the Lagrange multiplier and the dual parameter are updated.

We introduce here a novel variant of GoDec, dubbed R-GoDec, which manages at the same time both the presence of outliers and unspecified entries in the data matrix $\widehat{X}$. More in detail, the sparse term is expressed as the sum of two terms S_1 and S_2 having complementary supports:

- S_1 is a sparse matrix with support on Ω representing outlier measurements;
- S_2 has support on $\mho$ and it is an approximation of $-\mathcal{P}_{\mho}(L)$, representing completion of missing entries.

This results in the following model

$$\mathcal{P}_{\Omega}(\widehat{X}) = L + S_1 + S_2 + N \tag{19}$$

which is the natural combination of the RPCA formulation (2) with the MC formulation (13) associated to the GoDec algorithm. Equation (19) reduces to (2) over Ω, since S_2 is zero in Ω. On $\mho$ instead, Equation (19) turns to $L + S_2 + N = 0$, since both S_1 and $\widehat{X}$ are zero in $\mho$, and thus S_2 must coincide with $-L$ (up to noise) as in the case of Problem (12).

The decomposition Problem (19) is translated into the following minimization

$$\begin{cases} \min\limits_{L,S_1,S_2} \dfrac{1}{2}\left\|\mathcal{P}_{\Omega}(\widehat{X}) - L - S_1 - S_2\right\|_F^2 + \lambda\,\|S_1\|_1 \\ \text{s.t. } \operatorname{rank}(L) \leqslant r, \\ \operatorname{supp}(S_1) \subseteq \Omega, \\ \operatorname{supp}(S_2) = \mho \end{cases} \tag{20}$$

which is solved using a block-coordinate scheme that alternates the steps of Algorithm 1 and Algorithm 2. First, the rank-r projection of $\mathcal{P}_{\Omega}(\widehat{X}) - S_1 - S_2$ (computed through BRP) is assigned to L. Then, the sparse terms S_1 and S_2 are updated separately. The outlier term S_1 is computed by applying the soft-thresholding operator Θ_{λ} to the matrix $\mathcal{P}_{\Omega}(\widehat{X} - L)$. As for the completion term, $-\mathcal{P}_{\mho}(L)$ is assigned to S_2, according to the GoDec algorithm for MC. These steps are iterated until convergence. Our method, called R-GoDec where "R" stands for "robust", is summarized in Algorithm 3. Note that R-GoDec has a flexible structure, *e.g.*, it can be extended to the case where $\mathcal{P}_{\Omega}(\widehat{X})$ contains Euclidean distances (Rossi et al., 2017).

The computational complexity of R-GoDec is dominated by the computation of L via BRP, which requires, as reported in (Zhou and Tao, 2011), $r^2(n_1 + 3n_2 + 4r) + (4q+4)n_1 n_2 r$ flops, where $\widehat{X}$ is $n_1 \times n_2$ dense matrix with rank r and q is an integer (between 1 and 10 in our experiments) that controls the approximation error of the BRP. Step 2 requires $n_1 n_2$ flops and the test of the cycle takes $5n_1 n_2$ flops. In total R-GoDec requires $(r\,(4\,q+4)+6)\;n_1\,n_2 + r^2\,n_1 + \left(3\,r^2\right)\;n_2 + 4\,r^3$ flops per iteration, that asymptotically is $O(n_1 n_2)$.

Algorithm 3 R-GoDec

Input: $\widehat{X}$, Ω, r, ϵ, λ
Output: L, S_1, S_2
Initialize: $L = \widehat{X}$, $S_1 = 0$, $S_2 = 0$
while $\left\|\mathcal{P}_{\Omega}(\widehat{X}) - L - S_1 - S_2\right\|_F^2 / \left\|\mathcal{P}_{\Omega}(\widehat{X})\right\|_F^2 > \epsilon$ **do**
1. $L \leftarrow$ rank-r approximation of $\mathcal{P}_{\Omega}(\widehat{X}) - S_1 - S_2$ via BRP
2. $S_1 \leftarrow \Theta_{\lambda}(\mathcal{P}_{\Omega}(\widehat{X} - L))$
3. $S_2 \leftarrow -\mathcal{P}_{\mho}(L)$

end while

We shall summarize here what we know about the convergence of R-GoDec, discussing first the convergence of the objective function and then the convergence of the variables. In this analysis we assume that the matrix L is produced by SVD instead of BRP, i.e., we disregard the random nature of BRP and assume a deterministic result.

Convergence of the objective function. The alternative minimization of problem (20) with respect to one variable at a time produces a convergent sequence of objective function values $\frac{1}{2}\left\|\mathcal{P}_{\Omega}(\widehat{X}) - L - S_1 - S_2\right\|_F^2 + \lambda\,\|S_1\|_1$. Indeed, each of the three sub-problems has an optimal solution that can be computed respectively by updating L via singular value hard-thresholding of $\mathcal{P}_{\Omega}(\widehat{X}) - S_1 - S_2$, S_1 via entry-wise soft-thresholding of $\mathcal{P}_{\Omega}(\widehat{X} - L)$, and S_2 via $-\mathcal{P}_{\Omega}(L)$. Therefore, in each step the sub-problem over the coordinate block is solved exactly to its optimal solution hence the sequence of non-negative objective function values is non-increasing.

Convergence of the variables L, S_1 and S_2. Convergence results for GoDec are known in the literature in the two following cases:

Robust PCA (no missing data). In this case S_2 is null and S_1 is not null; then the minimization problem (20) reduces to the RPCA case, namely problem (5) where $S = S_1$. In this case the cost function is continuous and coercive, in fact, intuitively, if S grows to infinity so does the cost, while if L grows to infinity then S seeks to compensate the growth pushing $||S||_1$ and thus the cost to infinity. Thus there exists a convergent subsequence of variables. Thanks to Lemma 3.1 and Theorem 4.1 in (Tseng, 2001) the cluster point of such subsequence is actually a coordinate-wise minimum and stationary point. The convergence of the low-rank term is reported in (Zhou and Tao, 2013).

Matrix completion (no outliers). In this case S_1 is null and S_2 is not null; then the minimization problem (20) collapses to the MC case, namely problem (10) where $S = S_2$. In this case the linear convergence of the variables to a local minimum can by proved within the framework of *alternating projections on two manifolds* (Lewis and Malick, 2008), as stated in Theorem 3 in (Zhou and Tao, 2011).

R-GoDec instead is a mix of these two cases, since it caters for outliers *and* missing entries. In this case S_1 and S_2 are both

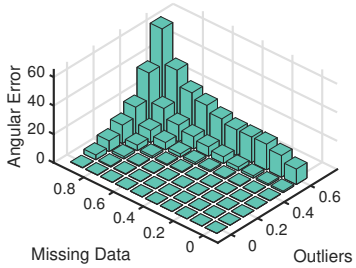


Figure 4: Angular error [degrees] achieved by R-GoDec on rotation synchronization versus fraction of outliers and missing data. The two cases of MC and RPCA correspond to the axes.

present, and we have not been able to prove convergence results for the variables. In particular, the cost function (20) is no longer coercive, and also the problem does not fall any more under the framework of *alternating projections on two manifolds*. Nevertheless, empirical evidence supports the conjecture that there is a region which extends beyond the axes where R-GoDec converges to the global minimum of (20). Figure 4 reports the result of a simulation where we varied the percentage of outliers and the percentage of missing entries, as described in Section 7.1. Since no noise has been added, zero error means that the correct solution is attained. This occurs on a region around the origin, which defines empirically the convergence basin of R-GoDec.

Please note that even on the axis the error is not always zero: in fact it increases when outliers reach the breakdown point or missing entries approaches the limit of matrix completion. For this reason, in the mixed case, when the error departs from zero this can be caused by non-convergence to a global minimum or by the intrinsic limits of the method itself, but we cannot tell which alternative applies.

5. Problem formulation

The *synchronization* problem (Singer, 2011) aims at recovering elements of a group given measures of their ratios, thus leaping from two-view to multi-view information. More precisely, let Σ denote a group, let $Y_1, \ldots, Y_n$ denote n elements of Σ and let $X_{ij} \in \Sigma$ denote the ratio between Y_i and Y_j. The goal is to recover $Y_1, \ldots, Y_n \in \Sigma$ based on the knowledge of a certain number of $X_{ij} \in \Sigma$ such that the following *compatibility constraint* is satisfied

$$X_{ij} = Y_i Y_j^{-1}. \tag{21}$$

In this paper we are interested in $\Sigma = SE(3)$, in which case we are in the context of *rigid-motion synchronization* (also known as *motion averaging* (Govindu, 2004)), and in $\Sigma = SO(3)$, in which case we are dealing with *rotation synchronization* (also known as *rotation averaging* (Hartley et al., 2013)). In the first case, $Y_1, \ldots, Y_n \in \mathrm{SE}(3)$ denote n rigid transformations representing the *absolute* – *i.e.*, expressed in an external coordinate system – angular attitudes and positions of local reference frames, and $X_{ij} \in \mathrm{SE}(3)$ denotes the ideal (noise-free) *relative* motion of the pair (i, j), namely the rigid transformation (or direct isometry) that maps the reference frame i in that associated to j. In the second case, $Y_1, \ldots, Y_n \in \mathrm{SO}(3)$ denote n rotations representing the *absolute* angular attitudes of local reference frames, and $X_{ij} \in \mathrm{SO}(3)$ denotes the ideal *relative* rotation of the pair (i, j).

In this paper we consider the matrix representations of SE(3) and SO(3), which are viewed as subgroups of the General Linear Group of degree 4 and 3, respectively

$$\mathrm{SE}(3) = \left\{ \begin{pmatrix} R & \mathbf{t} \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix}, \text{ s.t. } R \in \mathrm{SO}(3),\ \mathbf{t} \in \mathbb{R}^3 \right\} \tag{22}$$

$$\mathrm{SO}(3) = \{R \in \mathbb{R}^{3\times 3}, \text{ s.t. } R^\mathsf{T} R = R R^\mathsf{T} = I,\ \det(R) = 1\} \tag{23}$$

thus inverse and composition of rotations/rigid transformations reduce to matrix operations. In particular, in the Special Orthogonal Group the inverse equals matrix transposition, thus the compatibility constraint rewrites

$$X_{ij} = Y_i Y_j^\mathsf{T}. \tag{24}$$

In practice only a subset of all the relative motions/rotations is available, due to the lack of overlap between some pairs of images/scans. However, there is a significant level of redundancy in general datasets, which can be used to distribute the error over all the nodes, avoiding drift in the solution. Let A be the $(0,1)$-matrix that indicates the available measurements: $A_{ij} = 1$ if X_{ij} is available, $A_{ij} = 0$ otherwise. Let us consider the *measurement graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ whose adjacency matrix is A, where $\mathcal{V}$ denotes the vertex set and $\mathcal{E}$ denotes the edge set. It must consist of a single connected component, in order to guarantee solvability of the synchronization problem. Consequently, in the minimal case the graph must be a spanning tree over n nodes, which has $n - 1$ edges (*i.e.*, relative motions/rotations).

Let $\widehat{X}_{ij}$ denote an estimate of X_{ij} (in this paper we use the hat accent to denote noisy measurements). The estimated relative rotations/motions are usually corrupted by a diffuse noise with small variance, hence they do not satisfy Equation (21) exactly. Thus the goal is to average them so as to maximally satisfy $\widehat{X}_{ij} \approx Y_i Y_j^{-1}$. A possible approach consists in formulating the following optimization problem

$$\min_{Y_i \in \Sigma} \sum_{(i,j)\in\mathcal{E}} \left\| \widehat{X}_{ij} - Y_i Y_j^{-1} \right\|_F^2 \tag{25}$$

where the Frobenius norm $\|\cdot\|_F$ defines a left-invariant metric on SE(3) and a bi-invariant metric on SO(3). The solution is determined up to a global transformation, affecting the external coordinate system. This fact is inherent to the problem and cannot be resolved without external measurements. Note that if $\mathcal{G}$ is a tree then there is no counteraction of the errors in the solution. However, as soon as redundant measures are considered (*i.e.* the graph has at least one cycle), they are exploited by the synchronization process to globally compensate the errors.

Problem (25) is analyzed in depth in (Hartley et al., 2013) in the $\Sigma = SO(3)$ case under the name ℓ_2-*chordal averaging*. Further theoretical analysis is reported in (Wilson et al., 2016) where it is shown that smaller and well-connected graphs are

easier than larger and noisy ones, based on a local convexity analysis.

As observed in (Arie-Nachimson et al., 2012), Problem (25) can be reformulated in a useful equivalent form that captures a low-rank constraint.

Let X denote the $rn \times rn$ block-matrix containing the relative measures, let Y denote the $rn \times r$ block-matrix containing the absolute transformations, and let $Y^{-\flat}$ denote the $r \times rn$ block-matrix containing the inverse of absolute transformations, namely

$$X = \begin{pmatrix} I & X_{12} & \dots & X_{1n} \\ X_{21} & I & \dots & X_{2n} \\ \dots & & & \dots \\ X_{n1} & X_{n2} & \dots & I \end{pmatrix}, \quad Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix},$$

$$Y^{-\flat} = \begin{bmatrix} Y_1^{-1} & Y_2^{-1} & \dots & Y_n^{-1} \end{bmatrix} \tag{26}$$

where I denotes the identity matrix, $r = 3$ in the $\Sigma = \mathrm{SO}(3)$ case and $r = 4$ in the $\Sigma = \mathrm{SE}(3)$ case. Using this notation, the compatibility constraint (21) can be expressed in a compact form as

$$X = YY^{-\flat} \tag{27}$$

which implies that $\mathrm{rank}(X) = r$. In the Special Orthogonal Group Equation (27) reduces to

$$X = YY^T \tag{28}$$

and hence, besides being low-rank, the matrix X is symmetric and positive semidefinite.

The minimum number of relative measures necessary to solve the synchronization problem is $n-1$; in addition to this, we have n measures equal to the identity matrix along the diagonal of X (*i.e.*, $X_{ii} = I$ for $i = 1, \dots, n$) resulting in $r^2(2n-1)$ specified entries of X. Not surprisingly, this number coincides with the degrees of freedom of a generic $rn \times rn$ rank-r matrix (in general, a $n_1 \times n_2$ matrix of rank r has $(n_1 + n_2 - r)r$ degrees of freedom).

Let $\widehat{X}$ be a noisy version of the ideal matrix X containing the observed relative rotations/motions $\widehat{X}_{ij}$. Since the measurement graph is not complete in general, $\widehat{X}$ has missing entries, which are represented as zero blocks. In other words, the available relative information is given by $\mathcal{P}_\Omega(\widehat{X})$, where the sampling set has a $r \times r$ block-structure: $\Omega = A \otimes \mathbb{1}_{r\times r}$, where $\otimes$ denotes the Kronecker product and $\mathbb{1}_{r\times r}$ is a $r \times r$ matrix of ones. Using this notation, the minimization problem (25) can be expressed as

$$\begin{aligned} &\min_X \ \left\| \mathcal{P}_\Omega(\widehat{X} - X) \right\|_F^2 \\ &\text{s.t. } X = YY^{-\flat},\ Y \in \Sigma^n. \end{aligned} \tag{29}$$

Problem (29) is non-convex, and therefore computationally hard to solve in general. For this reason, a principled strategy consists in formulating tractable approaches that solve the problem approximately but accurately, so as to provide a good starting point for a subsequent local refinement. This can be regarded as an "extrinsic calculation", for the rotation/rigid-motion constraints are relaxed to compute the solution.

One example is the *spectral relaxation* where the synchronization problem is cast to a spectral decomposition, and a straightforward robust enhancement can be gained via Iteratively Reweighted Least Squares (IRLS). This approach was introduced by Singer (2011) for SO(2), extended in (Singer and Shkolnisky, 2011; Arie-Nachimson et al., 2012) to SO(3) and further generalized in (Arrigoni et al., 2016c) to SE(3). Another possibility is the *semidefinite relaxation* used in (Arie-Nachimson et al., 2012; Rosen et al., 2015) where the synchronization problem is expressed as a semidefinite program (SDP).

6. Proposed approach

In this section we cast the synchronization problem as a LRS matrix decomposition, paving the way to the application of general matrix decomposition techniques in structure-from-motion and multiple 3D point-set registration. Besides this, we adapt R-GoDec (Algorithm 3) in order to fit the needs of rotation and rigid-motion synchronization.

We observe that the formulation in (29) can successfully average noisy relative measures, but it is not resistant to outliers. For this reason we consider the following problem

$$\begin{aligned} &\min_{X,S} \ \left\| \mathcal{P}_\Omega(\widehat{X} - X) - S \right\|_F^2 \\ &\text{s.t. } X = YY^{-\flat},\ Y \in \Sigma^n,\ S \text{ is sparse in } \Omega \end{aligned} \tag{30}$$

where the additional variable S represents outliers, which are sparse over the measurement graph (by assumption).

If the *rank relaxation* is adopted, *i.e.* all the constraints except of the rank property are ignored, then the following optimization problem is obtained

$$\begin{aligned} &\min_{L,S} \ \left\| \mathcal{P}_\Omega(\widehat{X} - L) - S \right\|_F^2 \\ &\text{s.t. } \mathrm{rank}(L) \leqslant r,\ S \text{ is sparse in } \Omega \end{aligned} \tag{31}$$

where L denotes a low-rank matrix which approximates the theoretical X defined in (26). It is worth noting that an outlier term is included in the cost function by design. With respect to non robust solutions that rely on a preliminary outlier rejection step, this approach has the great advantage of being *intrinsically* resilient against outliers.

Note that Problem (31) is indeed a LRS decomposition problem with unspecified entries and outliers, since it is associated to the formulation (14), namely $\mathcal{P}_\Omega(\widehat{X}) = \mathcal{P}_\Omega(L) + S + N$. Thus the absolute rotations/motions can be recovered by means of any algorithm that computes such decomposition, such as Grasta (He et al., 2012) or L1-Alm (Zheng et al., 2012), as explained in Section 4.3. Alternatively, the equivalent formulation (19) can be used, namely $\mathcal{P}_\Omega(\widehat{X}) = L + S_1 + S_2 + N$, which can be computed via the R-GoDec algorithm.

In the synchronization problem, however, the data matrix $\widehat{X}$ has a block structure, being composed of rotations or rigid motions, and this should be reflected by the sparse term which represents the outliers. This is taken into account by modifying Algorithm 3 in order to enforce a block-structure in S_1. Specifically, the ℓ_1-norm in (20) is substituted with the mixed $\ell_{2,1}$-norm which promotes group sparsity. Accordingly, we address

the following problem

$$\begin{cases} \min\limits_{L,S_1,S_2} \frac{1}{2} \left\| \mathcal{P}_\Omega(\widehat{X}) - L - S_1 - S_2 \right\|_F^2 + \lambda \left\| S_1 \right\|_{2,1} \\ \text{s.t. } \operatorname{rank}(L) \leqslant r, \\ \operatorname{supp}(S_1) \subseteq \Omega, \\ \operatorname{supp}(S_2) = \mho \end{cases} \tag{32}$$

where the mixed $\ell_{2,1}$-norm of a $rn \times rn$ matrix S is defined as the sum of the Frobenius norm of each $r \times r$ block S_{ij}

$$\left\| S \right\|_{2,1} = \sum_{i,j=1}^{n} \left\| S_{ij} \right\|_F . \tag{33}$$

The minimum of the cost function with respect to S_1 keeping the other variables constant has a closed-form expression, given by the *generalized soft-thresholding* (or shrinkage) operator $\Theta_\lambda^{2,1}$ applied to the matrix $\mathcal{P}_\Omega(\widehat{X} - L)$ (Kowalski and Torrésani, 2009). Such an operator takes a $rn \times rn$ matrix S as input and on each $r \times r$ block S_{ij} it computes

$$\Theta_\lambda^{2,1}(S_{ij}) = S \cdot \max(1 - \frac{\lambda}{\left\| S_{ij} \right\|_F}, 0) \tag{34}$$

where scalar operations are applied element-wise. In this way the selected blocks are the ones with the biggest Frobenius norms. Accordingly, Step 2 of Algorithm 3 is modified as follows

$$S_1 \quad \leftarrow \Theta_\lambda^{2,1}(\mathcal{P}_\Omega(\widehat{X} - L)).$$

The optimal S_1 can be used to identify the outliers, since rogue relative measures correspond to nonzero blocks in S_1. Please note that, as in Algorithm 3, the cost function in (32) has a unique minimum with respect to S_1.

Once the optimal L is found via any LRS algorithm, we proceed as follows to compute the absolute rotations/motions. Since the solution to synchronization is defined up to a global transformation, any block-column of L can be used as an estimate of Y. Due to the rank relaxation, each $r \times r$ block is not guaranteed to belong to Σ and needs to be projected onto the group. As concerns the Special Orthogonal Group, the nearest rotation matrix (in the Frobenius norm sense) is found via Singular Value Decomposition (SVD) (Keller, 1975). As concerns the Special Euclidean Group, every fourth row is set equal to $[0\ 0\ 0\ 1]$ and 3×3 rotation blocks are projected onto SO(3) through SVD, as explained in (Belta and Kumar, 2002).

7. Experiments in SO(3)

We plugged R-GoDec, Grasta and L1-Alm in our framework, obtaining three rotation synchronization methods based on LRS matrix decomposition. We evaluated these solutions on both synthetic and real scenarios in terms of accuracy, execution cost and robustness to outliers. All the experiments were performed in Matlab on a dual-core MacBook Air with i5 1.3GHz processor, 4Gb RAM.

We compared R-GoDec, Grasta and L1-Alm to several techniques from the state of the art. We considered the spectral relaxation (EIG) (Arie-Nachimson et al., 2012) and its robust variation (EIG-IRLS), the semidefinite relaxation (SDP) (Arie-Nachimson et al., 2012), the OptSpace algorithm (Keshavan et al., 2010), the Weiszfeld algorithm (Hartley et al., 2011), the L1-IRLS algorithm (Chatterjee and Govindu, 2013), and the LUD algorithm (Wang and Singer, 2013).

The code of LUD has been provided by Wang and Singer (2013), the codes of Grasta, L1-Alm, OptSpace and L1-IRLS are available on the web, while in the other cases we used our implementation. The Cauchy weight function (Holland and Welsch, 1977) was used in EIG-IRLS and the SeDuMi toolbox (Sturm, 1999) was used to solve the semidefinite program associated to the SDP method. Our implementation of R-GoDec is available on-line[1]. All the methods used the default tuning parameter(s) specified in the original paper or code.

In order to compare estimated and ground-truth absolute rotations we employed ℓ_1 single averaging. Specifically, if $\widehat{Y}_1, \ldots, \widehat{Y}_n$ are estimates of the theoretical absolute rotations $Y_1, \ldots, Y_n$, then the optimal $S \in SO(3)$ that aligns them into a common reference system solves $Y_i = \widehat{Y}_i S$, and hence it is the single mean of the set $\{Y_i \widehat{Y}_i^T, i = 1, \ldots, n\}$, and it can be computed *e.g.* by using (Hartley et al., 2011). Then we used the angular distance to evaluate the accuracy of rotation recovery. The angular (or geodesic) distance between two rotations A and B is the angle of the rotation BA^T (in the angle-axis representation) so chosen to lie in the range $[0, 180°]$, namely $d_\angle(A, B) = d_\angle(BA^T, I) = 1/\sqrt{2} \left\| \log(BA^T) \right\|_2$. Other distances in SO(3) can be considered with comparable results.

7.1. Simulated Data

In our simulations we considered n rotation matrices sampled from random Euler angles, representing ground truth absolute rotations. The measurement graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a random graph drawn from the Erdős-Rényi model with parameters (n, p), *i.e.* given a vertex set $\mathcal{V} = \{1, 2, \ldots, n\}$ each edge (i, j) is in the set $\mathcal{E}$ with probability $p \in [0, 1]$, independently of all other edges. Thus $(1 - p)$ controls the degree of sparsity of the graph and $p = 1$ corresponds to the complete graph. Only connected graphs are considered among all the instances generated in this way. A fraction of the pairwise rotations was drawn uniformly from SO(3), simulating outliers. The remaining pairwise rotations were corrupted by multiplicative noise $\widehat{X}_{ij} = X_{ij} N_{ij}$ where $N_{ij} \in$ SO(3) has axis uniformly distributed over the unit sphere and angle following a Gaussian distribution with zero mean and standard deviation $\sigma_R \in [1°, 10°]$, thus representing a small perturbation of the identity matrix. Considering the first order approximation of rotations, this corresponds to additive noise. All the results were averaged over 50 trials.

It is hard to evaluate the performances of a synchronization method as a whole, since several factors are involved, thus in the following simulations we let one parameter vary at a time and keep the others fixed.

[1] www.diegm.uniud.it/fusiello/demo/gmf/

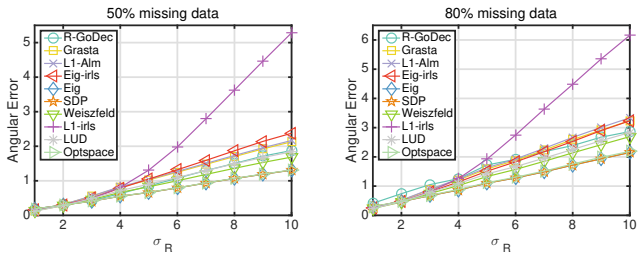


Figure 5: Mean angular errors [degrees] as a function of the noise standard deviation, with $p = 0.5$ (left) and $p = 0.2$ (right). Outliers are not introduced in this experiment.

Noise

In this experiment we analyse the behaviour of the aforementioned methods in the presence of noise among the input rotations without introducing outliers, with $n = 100$. Results are reported in Figure 5 with $p = 0.5$ and $p = 0.2$, which correspond to about 50% and 80% of missing pairs, respectively. As expected the lowest errors are achieved by the non-robust methods, namely EIG, SDP and OptSpace. On the contrary, all the robust methods yield worse results, since they essentially trade robustness for statistical efficiency. This compromise is particularly evident in L1-IRLS at the point when it switches from quadratic to fixed loss, defined by a fixed value (5°), whereas EIG-IRLS, that uses a data dependent threshold, has a more linear trend, for the trade-off takes place at all noise levels.

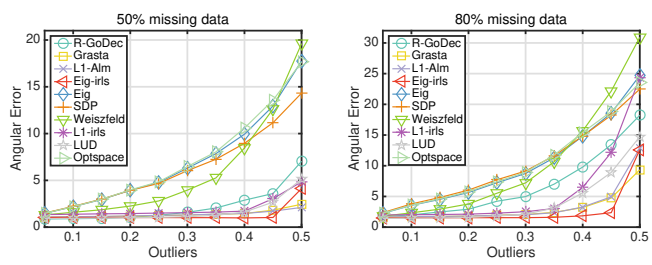


Figure 6: Mean angular errors [degrees] as a function of q, with $p = 0.5$ (left) and $p = 0.2$ (right). A fixed level of noise is applied to the inlier rotations in this experiment.

Outliers

In this experiment we study the robustness to outliers of our approach. Each edge $(i, j) \in \mathcal{E}$ was designated as an outlier with uniform probability $q \in [0, 1]$, independently of all other edges. Figure 6 shows the angular errors of all the analysed methods as a function of q, with $n = 100$. As before, we chose $p = 0.2$ and $p = 0.5$ to define the density of the measurement graph. In this experiment all the inlier rotations were corrupted by a fixed level of noise ($\sigma_R = 5°$). When the percentage of unspecified relative rotations is about 50%, the errors of R-GoDec, Grasta and L1-Alm remain almost constant, showing no sensitivity to outliers. The same happens for LUD, L1-IRLS and EIG-IRLS. On the contrary, EIG, SDP and OptSpace are not robust to outliers, as already observed in the previous sections. As for the Weiszfeld algorithm, its performances places it at the middle between robust and non-robust solutions. Specifically, it shows good resilience to outlier rotations when they are below 30%, then the errors start to grow up, yielding a behaviour similar to non-robust approaches. When the data matrix is highly incomplete ($p = 0.2$), the difference between robust and non-robust solutions becomes smaller, however results are qualitatively similar to the previous case.

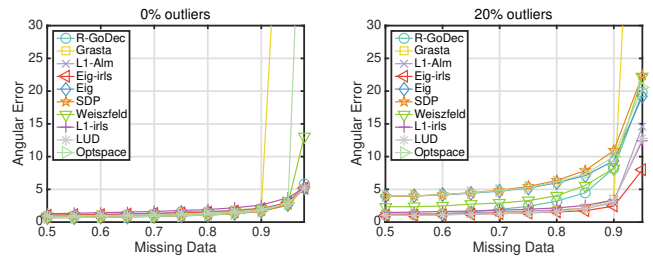


Figure 7: Mean angular errors [degrees] as a function of $(1 - p)$, with $q = 0$ (left) and $q = 0.2$ (right). A fixed level of noise is applied to the inlier rotations in this experiment. The average angular error of Grasta is approximately 80° for $(1 - p) = 0.95$.

Missing Data

In this experiment we study how missing data influence the performances of LRS algorithms. Figure 7 reports the angular errors of the analysed methods as a function of $(1 - p)$, with $n = 100$. The sparsity parameter $(1 - p)$ ranges from 0.5 to 0.95, which correspond to about 50% and 95% of missing pairs. Results with lower values of $(1-p)$ yield the same behaviour as $(1 - p) = 0.5$, and hence they are not reported. We considered both the ideal case where outliers are not present ($q = 0$) and a more realistic situation in which a given percentage of outliers is introduced ($q = 0.2$). In the first case we also considered the minimal situation in which $n - 1$ relative rotations are available, which corresponds to 98% of missing pairs. In both cases all the inlier rotations were corrupted by a fixed level of noise ($\sigma_R = 5°$). In the absence of outliers, when the percentage of missing pairs do not exceed 90%, the errors obtained by our approach and the remaining methods remain constant, showing no sensitivity to missing data. Grasta can tolerate up to 90% of missing pairs, whereas R-GoDec and L1-Alm can also handle the minimal situation, but the errors are higher then the previous cases. Indeed, there is no way to compensate the initial errors since there is no redundancy. However, it should be noted that if only $n - 1$ relative rotations are available, then there is no need to perform rotation synchronization, and the absolute rotations can be computed by propagating the compatibility constraint (21) along a spanning tree, starting from any node assumed equal to the identity matrix.

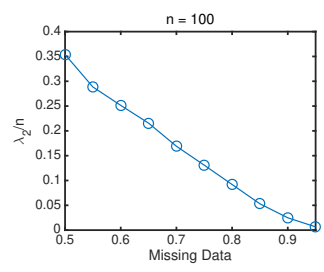


Figure 8: Algebraic connectivity of $\mathcal{G}$ divided by the number of absolute rotations, as a function of $(1-p)$. Low values correspond to hard problems (Wilson et al., 2016).

In the presence of outliers, Grasta can tolerate up to 90% of missing data, while R-GoDec and L1-Alm give reasonable results until 95% of missing data. However, the performances of R-GoDec degrade starting from 80% of missing pairs, reaching errors comparable to non-robust methods when the percentage of missing pairs is 90%. As for the remaining algorithms, their behaviour is qualitatively similar to the case where outliers are not present.

In order to provide a quantitative measure of hardness of the synchronization problems related to Figure 7, we consider the indicator introduced by Wilson et al. (2016)

$$\frac{\lambda_2(\mathcal{G})}{n} \tag{35}$$

where $\lambda_2(\mathcal{G})$ denotes the algebraic connectivity of the measurement graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, that is the second-smallest eigenvalue of the Laplacian matrix. Figure 8 reports the value of $\lambda_2(\mathcal{G})/n$ as a function of the sparsity parameter $(1-p)$ with $n = 100$, where low values correspond to hard problems, according to the analysis in (Wilson et al., 2016). As expected, rotation synchronization becomes more difficult to solve as the percentage of missing data increases.

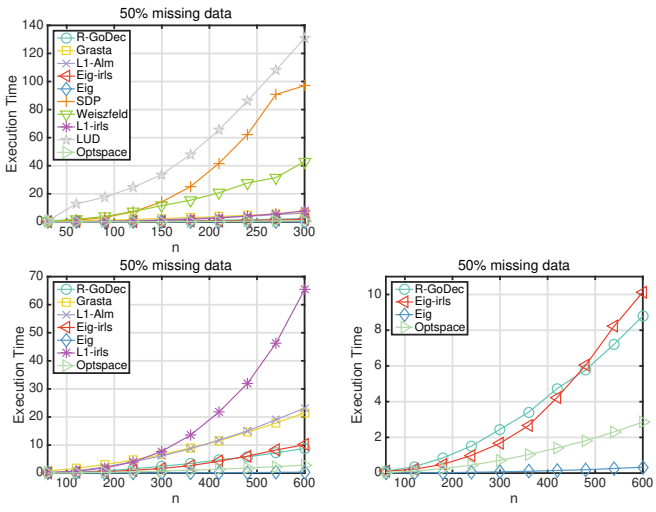


Figure 9: Execution times [seconds] as a function of the number of absolute rotations, with $p = 0.5$ and $q = 0.2$. A fixed level of noise is applied to the inlier rotations in this experiment. LUD, SDP and Weiszfeld are analysed only with a maximum of 300 nodes due to computational limitations. The bottom row figures are a magnification of the top-left one.

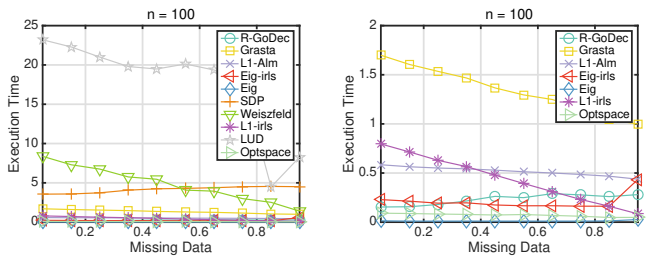


Figure 10: Execution times [seconds] as a function of $(1-p)$, with $n = 100$ and $q = 0.2$. A fixed level of noise is applied to the inlier rotations in this experiment. The right figure is a magnification of the left one.

Execution Time

In this experiment we analyse the computational efficiency of our approach in two situations. First, we kept the density level of the measurement graph fixed ($p = 0.5$) and let n vary between 30 and 600. Then, we kept the number of absolute rotations fixed ($n = 100$) and let $(1-p)$ vary between 0.05 (about 5% of missing data) and 0.95 (about 95% of missing data). In both cases we introduced a fixed level of noise and outliers on relative rotations ($\sigma_R = 5^\circ$, $q = 0.2$).

Figure 9 shows that LUD, SDP and Weiszfeld qualify as the slowest algorithms, while the other ones are significantly faster. In particular, the EIG method is the fastest solution to the rotation synchronization problem, but it is not robust. Among all the robust methods, R-GoDec and EIG-IRLS achieve the lowest execution times, outperforming L1-IRLS. The execution times of Grasta and L1-Alm are slightly higher than R-GoDec.

Figure 10 shows that LUD and Weiszfeld require more time as the viewing graph gets denser, whereas the execution times of the other techniques do not change significantly as p varies.

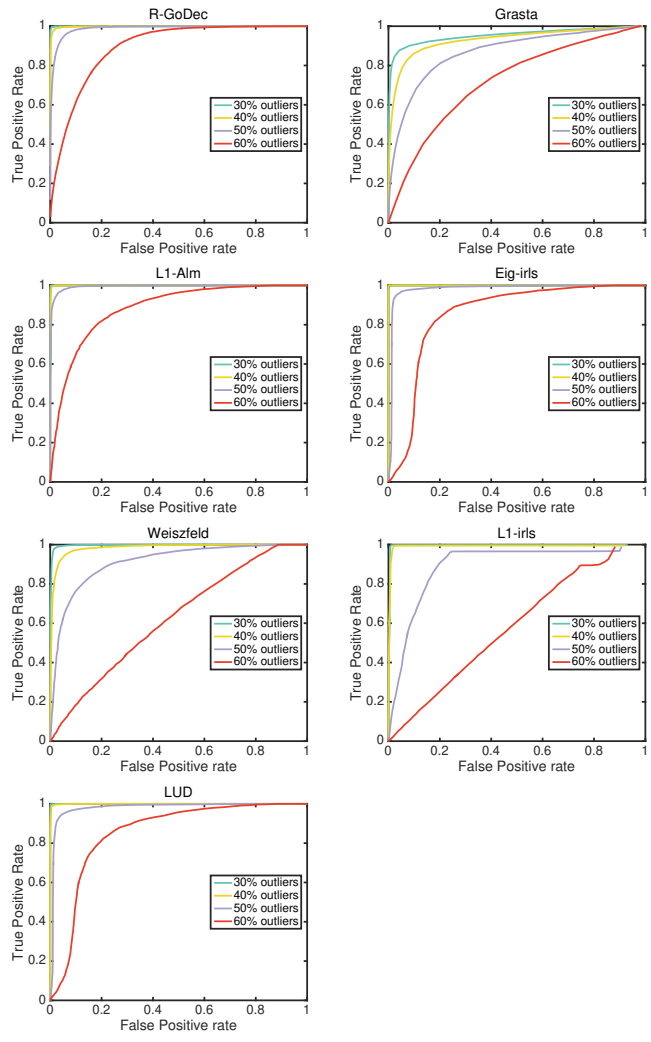


Figure 11: ROC curves of outlier detection for different percentages of outliers, with $p = 0.2$. A fixed level of noise is applied to the inlier rotations in this experiment.

Outlier Detection

We conclude this analysis by discussing the performances of our approach in terms of outlier detection, although this is not strictly part of rotation synchronization. Outliers correspond to non-zero entries in the sparse term, namely S_1 for R-GoDec and S for Grasta. We considered the receiver operating characteristic (ROC) curve, where the x-axis is the fraction of inliers erroneously classified as outliers, and the y-axis is the fraction of outliers correctly detected. The parameters that balance sparsity of outliers and noise are λ for R-GoDec and ρ for Grasta, thus each point in the ROC space is associated to a specific value of λ or ρ, respectively.

While R-GoDec and Grasta separate outliers and noise, the L1-Alm algorithm does not perform a classification of the data into inliers/outliers, thus we used a threshold on the error residuals. More precisely, a relative rotation $\widehat{X}_{ij} \in \mathrm{SO}(3)$ is classified as inlier if

$$d_{\angle}(\widehat{X}_{ij}, \widehat{Y}_i \widehat{Y}_j^{\mathsf{T}}) \leqslant \theta \tag{36}$$

where $\widehat{Y}_1, \ldots, \widehat{Y}_n$ are estimates of the absolute rotations, thus each point in the ROC space is associated to a specific value of θ. The same procedure was used also for EIG-IRLS, L1-IRLS, LUD and the Weiszfeld algorithm, whereas non robust methods are not considered in this experiment.

Figure 11 shows the ROC curves of the analysed algorithms for different percentages of outliers, with $p = 0.2$ and $n = 100$. All the inlier rotations were corrupted by a fixed level of noise ($\sigma_R = 5^\circ$). It is remarkable that R-GoDec gives a perfect classification with up to 50% of outliers. The performances drop with 60% of outliers, which is however a fairly high degree of contamination. Figure 11 evidences that R-GoDec is generally more accurate than Grasta in terms of classification, probably thanks to the mixed $\ell_{2,1}$-norm that promotes a block structure in S_1, whereas there are no significative differences between R-GoDec and L1-Alm. The ROC curves obtained by EIG-IRLS and LUD are comparable to those of R-GoDec and L1-Alm with up to 50% of outliers, and they are generally better than L1-IRLS and the Weiszfeld algorithm.

7.2. Structure-from-motion

We applied R-GoDec, Grasta and L1-Alm to the rotation synchronization stage of the structure-from-motion problem, considering both the benchmark set up by Strecha et al. (2008) and the irregular large-scale image collection assembled by Wilson and Snavely (2014).

EPFL Benchmark

The EPFL benchmark (Strecha et al., 2008) contain from 8 to 30 images and it provides ground-truth absolute rotations, which were used to evaluate the performances of the analysed methods. We followed a common structure-from-motion pipeline to obtain estimates of relative rotations. First, reliable matching points across the input images were computed by extracting and matching SIFT features (Lowe, 2004). Then, for each image pair, the essential matrix was computed in a RANSAC scheme, and it was factorized to obtain a unique $\widehat{X}_{ij} \in \mathrm{SO}(3)$ via SVD, which was considered missing if insufficient inlier correspondences were found. Relative rotations were then improved by applying bundle adjustment to pairs of cameras.

Results are shown in Table 1, which reports the mean angular errors and execution times of the analysed methods. R-GoDec, Grasta and L1-Alm are able to recover camera rotations accurately in low execution time, achieving better results than non-robust algorithms when contamination of outliers is particularly evident, namely in the Castle sequences which contain repetitive structures. In the remaining datasets all the analysed methods perform well, obtaining an average angular error less than 1°. Among the LRS methods, the highest accuracy is achieved by L1-Alm. Differences in execution times are meaningless for such relatively small datasets, except in the Castle-P30 sequence where LUD is remarkably slower than the other techniques.

Large-scale Datasets

The datasets from (Wilson and Snavely, 2014) contain from 247 to 2508 images and they provide estimates of relative rotations, which were given as input to all the analysed methods. These input rotations are very noisy and the associated graphs are highly incomplete in some cases, thus recovering camera rotations is challenging. Since ground-truth absolute rotations are not available, we used the output of Bundler (Snavely et al., 2006) as reference solution.

Results are shown in Table 2 and Figure 12, which report the median angular errors and the execution times of the analysed algorithms. Note that such errors are obtained *without* applying bundle adjustment, which is the final refinement required in any structure-from-motion system.

Neither EIG, SDP and OptSpace nor Weiszfeld and LUD are applicable in practical scenarios, since they do not satisfy the requirements of an efficient robust scheme. The first three achieve the highest errors since they are not robust to outliers, whereas the last two show resilience to outliers (to variable degrees) but they have the highest execution times.

The remaining algorithms solve the rotation synchronization problem while ensuring robustness and efficiency at the same time, to different extents. In particular, L1-IRLS achieves the highest accuracy, while R-GoDec and Grasta achieve an accuracy lower than L1-IRLS, albeit comparable in most datasets, in less time. L1-Alm is more accurate than R-GoDec and Grasta in most image sequences, at the expense of a higher execution time. If the percentage of missing data is extremely high, the performances of LRS methods degrades, in agreement with the outcome of the simulations and the literature on matrix completion. As for EIG-IRLS, results are comparable to L1-IRLS in most datasets both in terms of accuracy and execution time. R-GoDec and Grasta turn out to be the fastest solutions among all the robust ones.

Since there is an inherent trade-off between execution time and error in all the methods, the results reported in Figure 12 show that LRS methods are never dominated (in the Pareto sense) by others, meaning that they are candidates to lie on the Pareto front. In particular, R-GoDec reaches one limit of such a front, since it has the lowest execution time among all the robust algorithms. The opposite limit is represented by L1-IRLS that exhibits the lowest angular error.

Table 1: Mean angular errors [degrees] and execution times [seconds] for several algorithms on the datasets from (Strecha et al., 2008). The number of images (n), the percentage of missing relative rotations (miss) and the hardness measure in Equation (35) (λ_2/n) are also reported for each dataset. The lowest angular errors are highlighted in boldface.

				Grasta		R-GoDec		L1-Alm		L1-IRLS		EIG-IRLS		EIG		OptSpace		SDP		Weiszfeld		LUD	
Dataset	n	miss	λ_2/n	err	t	err	t	err	t	err	t	err	t	err	t	err	t	err	t	err	t	err	t
Castle-P30	30	55	0.07	0.55	0.35	1.33	0.04	0.38	0.08	0.65	0.04	**0.33**	0.32	2.03	0.01	2.06	0.06	2.22	0.40	0.89	0.87	0.56	10.13
Herz-Jesu-P25	25	43	0.08	0.68	0.35	0.10	0.02	0.09	0.06	**0.06**	0.03	**0.06**	0.11	0.13	0.01	0.14	0.05	0.14	0.30	**0.06**	0.93	0.07	0.97
Castle-P19	19	58	0.06	1.10	0.23	1.80	0.03	**0.80**	0.04	1.03	0.04	1.34	0.30	2.46	0.01	2.49	0.03	2.52	0.20	1.67	0.37	1.21	1.21
Fountain-P11	11	18	0.41	0.04	0.14	**0.03**	0.02	0.04	0.03	**0.03**	0.01	**0.03**	0.15	**0.03**	0.01	**0.03**	0.02	**0.03**	0.21	**0.03**	0.03	**0.03**	0.25
Entry-P10	10	2	0.80	0.07	0.13	0.43	0.01	**0.04**	0.03	0.30	0.04	0.05	0.11	0.52	0.01	0.52	0.01	0.52	0.16	0.25	0.10	0.07	1.14
Herz-Jesu-P8	8	18	0.44	0.05	0.09	0.05	0.02	**0.03**	0.06	0.04	0.02	0.04	0.18	0.05	0.01	0.05	0.02	0.05	0.16	0.05	0.02	0.06	0.15

Table 2: Median angular errors [degrees] and execution times [seconds] for several algorithms on the datasets from (Wilson and Snavely, 2014). The number of images (n), the percentage of missing relative rotations (miss) and the hardness measure in Equation (35) (λ_2/n) are also reported for each dataset. The lowest angular errors are highlighted in boldface.

				Grasta		R-GoDec		L1-Alm		L1-IRLS		EIG-IRLS		EIG		OptSpace		SDP		Weiszfeld		LUD	
Dataset	n	miss	λ_2/n	err	t	err	t	err	t	err	t	err	t	err	t	err	t	err	t	err	t	err	t
Piccadilly	2508	90	0.0002	6.41	72	15.31	41	11.47	387	**1.89**	468	23.95	424	34.99	21	18.65	33	-	-	16.81	1195	-	-
Roman Forum	1134	89	0.0008	2.65	21	13.08	5.6	8.95	75	2.27	17	**1.80**	45	21.46	2.1	14.57	5	-	-	22.54	250	-	-
Union Square	930	94	0.0004	6.87	12	9.25	2.1	14.63	46	**3.98**	9	4.84	29	8.62	0.8	10.15	3	10.98	3127	8.07	113	-	-
Vienna Cathedral	918	75	0.0006	2.08	20	3.11	3.3	1.83	52	**1.37**	64	1.60	37	5.96	2.2	5.43	5	6.15	2963	3.91	304	-	-
Alamo	627	50	0.0013	1.25	14	1.47	4.4	1.19	24	**1.09**	40	1.18	28	3.16	1.2	2.92	1.8	3.21	717	2.11	265	-	-
Notre Dame	553	32	0.0036	0.84	14	1.04	1.4	0.75	19	**0.65**	34	0.73	22	3.44	1.1	3.03	1.6	3.65	424	1.87	270	-	-
Tower of London	508	81	0.0017	2.77	6.2	3.36	0.7	2.90	13.4	**2.62**	2.9	2.78	6.4	3.86	0.3	3.72	0.9	3.98	380	3.31	73	-	-
Montreal N. Dame	474	53	0.0021	0.78	8.9	0.86	0.9	0.62	14	**0.57**	12	0.59	9.9	2.24	0.6	1.87	1.1	2.29	302	1.15	156	0.69	1038
Yorkminster	458	74	0.0016	2.04	5.9	2.31	3.0	1.85	11.4	**1.69**	3.3	1.82	6.5	5.84	0.3	4.97	0.9	5.67	280	3.75	80	1.87	937
Madrid Metropolis	394	69	0.0018	2.92	5	4.11	0.7	2.43	8.4	**1.01**	7.3	4.43	4	7.48	0.2	6.73	0.5	7.40	187	5.53	66	4.55	695
NYC Library	376	71	0.0025	2.26	4.7	3.40	0.3	1.76	7.8	**1.33**	3.2	1.99	3	5.51	0.2	5.28	0.6	5.58	149	3.68	59	1.95	681
Piazza del Popolo	354	60	0.0011	1.17	4.9	1.63	0.3	1.05	7.1	**0.98**	5.7	1.03	9.3	3.34	0.2	3.11	0.5	3.48	118	2.27	76	1.22	598
Ellis Island	247	33	0.0011	0.93	3.6	1.00	0.4	0.78	3.2	**0.57**	3.5	0.82	4.8	2.81	0.1	2.63	0.3	2.89	34	1.50	54	0.91	212

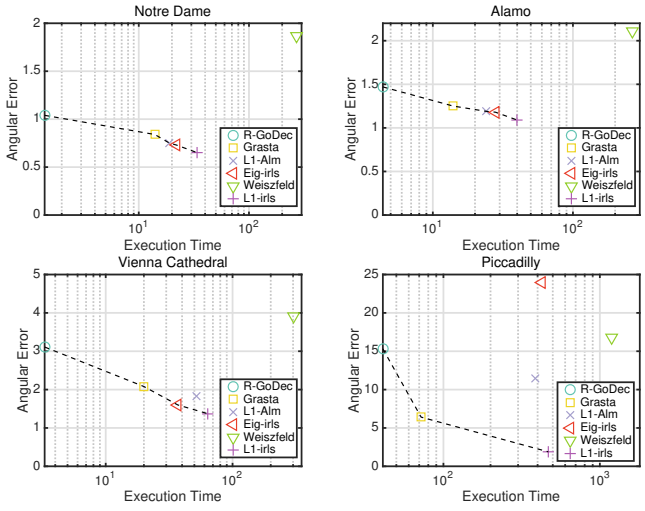


Figure 12: Median angular error [degrees] versus execution time [seconds]. The dashed line connects all the methods which are not dominated by any other method, meaning that no algorithm on this line has *both* execution time and angular error lower than the others. Only robust algorithms and the largest datasets from (Wilson and Snavely, 2014) are reported.

In summary, LRS methods achieve a good trade-off between robustness to outliers and speed. Failure cases appear when the percentage of missing data is extremely high.

8. Experiments in SE(3)

We evaluated our solution to synchronization in SE(3) on both simulated and real datasets, analysing resilience to noise, robustness to outliers, sensitivity to missing data and computational cost. We considered three LRS decomposition algorithms, namely R-GoDec, Grasta (He et al., 2012) and L1-Alm (Zheng et al., 2012), as done in Section 7.

We compared such algorithms to several techniques from the state of the art, namely the methods developed by Sharp et al. (2002), Govindu (2004), Torsello et al. (2011) (Diffusion), and Rosen et al. (2016). We also included in the comparison the spectral relaxation (EIG-SE(3)) and its robust variation (EIG-SE(3)-IRLS) (Arrigoni et al., 2016c). The codes of Grasta, L1-Alm, Diffusion and (Rosen et al., 2016) are available online, the one by Govindu (2004) has been provided by the author, while in the other cases we used our implementation. All the methods used the default tuning parameter(s) specified in the original paper or code.

In order to compare estimated and ground-truth absolute motions, we found the optimal isometry that aligns them by applying single averaging for the rotation term and least-squares for the translation term. Specifically, if $\widehat{Y}_1, \ldots, \widehat{Y}_n$ are estimates of the theoretical absolute motions $Y_1, \ldots, Y_n$ then the optimal $N \in \mathrm{SE}(3)$ that aligns them into a common reference system solves $Y_i = \widehat{Y}_i N$, which is equivalent to $R_i = \widehat{R}_i R$ and $\mathbf{t}_i = \widehat{R}_i \mathbf{t} + \widehat{\mathbf{t}}_\mathbf{i}$ by considering separately the rotation and translation term. Thus the optimal $R \in \mathrm{SO}(3)$ is the single mean of the set $\{R_i \widehat{R}_i^\mathsf{T}, i = 1, \ldots, n\}$, which can be estimated by applying ℓ_1 single averaging (Hartley et al., 2011), while the optimal $\mathbf{t} \in \mathbb{R}^3$ is computed in the least-squares sense. Then we used the angular distance and Euclidean norm to measure the accuracy of estimated rotations and translations respectively,

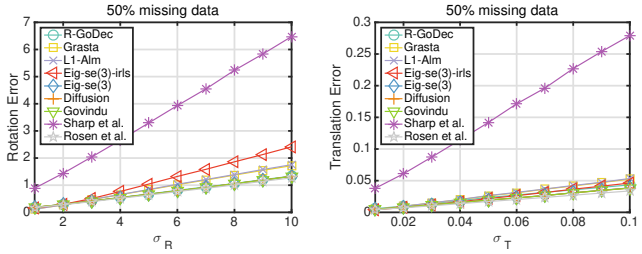


Figure 13: Mean errors on absolute motions as a function of the noise standard deviation, with $p = 0.5$. Outliers are not introduced in this experiment.

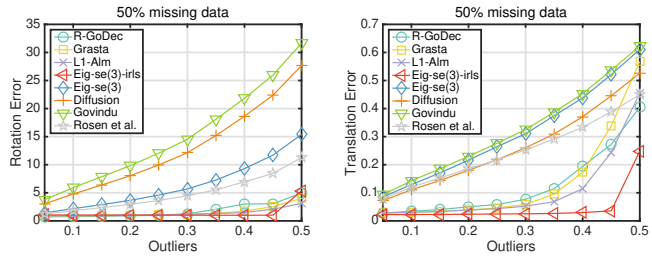


Figure 14: Mean errors on absolute motions as a function of q, with $p = 0.5$. A fixed level of noise is applied to the inlier relative motions in this experiment.

8.1. Simulated Data

We considered n absolute motions in which rotations were sampled from random Euler angles and translation coordinates followed a standard Gaussian distribution. As done in Section 7, the measurement graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ was drawn from the Erdős-Rényi distribution with parameters (n, p), and it was discarded if not connected. The inlier pairwise motions were corrupted by a multiplicative noise $\widehat{X}_{ij} = X_{ij}E_{ij}$, with $E_{ij} \in \mathrm{SE}(3)$ representing a small perturbation of the identity matrix. The rotation component of E_{ij} was generated as in Section 7, and the translation component was sampled from a Gaussian distribution with zero mean and standard deviation $\sigma_T \in [0.01, 0.1]$. All the results were averaged over 50 trials.

Noise

In this experiment we evaluate the effect of noise on relative motions in the absence of outliers. We considered $n = 100$ absolute motions and $p = 0.5$, which corresponds to about 50% of missing pairs. Figure 13 shows the mean errors on absolute motions (rotation errors are measured in degrees while translation errors are commensurate with the simulated data) obtained by all the analysed techniques, as a function of the standard deviation of noise. The worst resilience to noise is achieved by Sharp et al. (2002) while LRS decomposition techniques and the remaining algorithms return good estimates of absolute motions. A possible explanation of such behaviour is that in (Sharp et al., 2002) the error is distributed among the motions but it is not reduced. The best accuracy is achieved by non robust methods – namely, Diffusion, EIG-SE(3), Govindu (2004) and Rosen et al. (2016) – while robust techniques trade robustness for statistical efficiency.

Outliers

In this experiment we study the robustness to outliers of LRS methods. Each edge $(i, j) \in \mathcal{E}$ was designated as an outlier with uniform probability $q \in [0, 1]$, independently of all other edges. Outlier edges were assigned random elements of SE(3). We considered $n = 100$ absolute motions sampled as before, we chose $p = 0.5$ to define the density of the measurement graph, and we introduced a fixed level of noise on relative motions ($\sigma_T = 0.05, \sigma_R = 5^\circ$). The probability q that an edge is outlier ranges from 0.05 to 0.5, which correspond to about 5% and 50% of effective outliers. Results are reported in Figure 14, which shows the mean errors on absolute motions as a function of q. The errors obtained by Sharp et al. (2002) are not reported so as to better visualize differences between the remaining algorithms (the method by Sharp et al. (2002) yields an average rotation error of 20° for $q = 0.05$ and 100° for $q = 0.5$). Figure 14 confirms that Diffusion, EIG-SE(3) the methods by Govindu (2004) and Rosen et al. (2016) are not robust, and it clearly shows the resilience to outliers gained by R-GoDec, Grasta, L1-Alm and EIG-SE(3)-IRLS. In particular, the errors obtained by LRS decomposition techniques remain almost unchanged until $q = 0.4$ for rotations and $q = 0.3$ for translations.

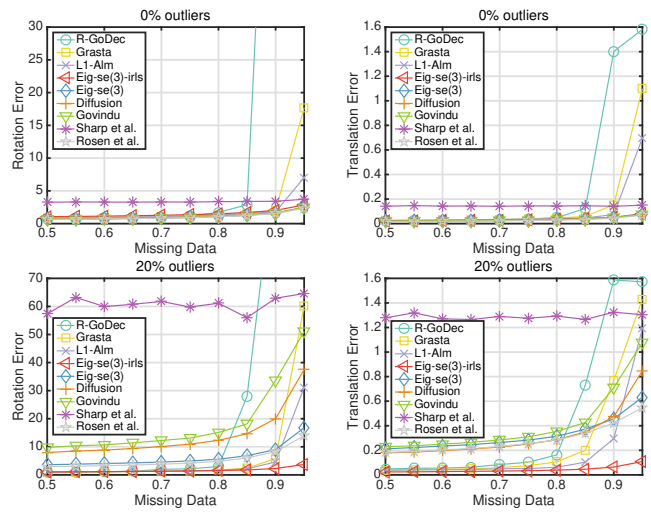


Figure 15: Mean errors on absolute motions as a function of $(1 - p)$, with $q = 0$ (top) and $q = 0.2$ (bottom). A fixed level of noise is applied to the inlier relative motions in this experiment. In the left sub-figures, the average rotation errors of R-GoDec are approximately 90° for $(1 - p) = 0.9$ and 120° for $(1 - p) = 0.95$.

Missing Data

In this experiment we study how missing data influence the performances of our approach. We considered $n = 100$ absolute motions sampled as before and we introduced a fixed level of noise on relative motions ($\sigma_T = 0.05, \sigma_R = 5^\circ$). The sparsity parameter $(1 - p)$ ranges from 0.5 to 0.95, which correspond to about 50% and 95% of missing pairs. We considered both the $q = 0$ case (no outliers) and the $q = 0.2$ case. Results are reported in Figure 15, which shows the mean errors on absolute motions as a function of the sparsity parameter $(1 - p)$. The errors obtained by Sharp et al. (2002) remain constant as $(1 - p)$ increases, showing no sensitivity to missing data. The same holds for the method by Govindu (2004), Rosen et al. (2016), Diffusion and the spectral relaxation, if there are

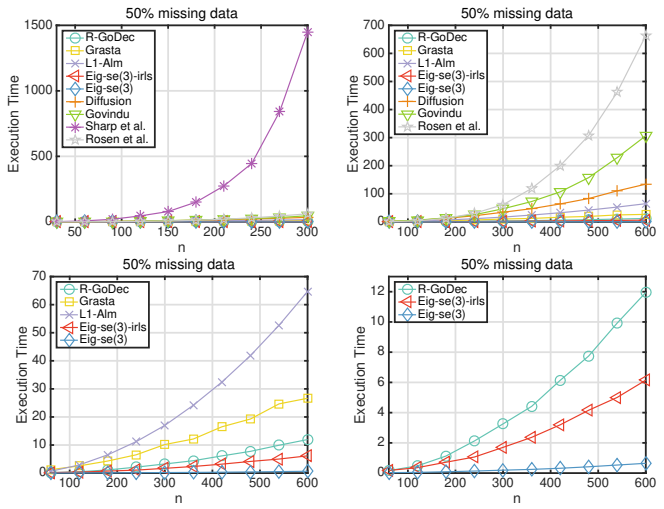


Figure 16: Execution times [seconds] as a function of n, with $p = 0.5$ and $q = 0.2$. A fixed level of noise is applied to the inlier relative motions in this experiment. The method by Sharp et al. (2002) is analysed only with a maximum of 300 nodes due to computational limitations. The top-right and bottom figures are a magnification of the top-left one.

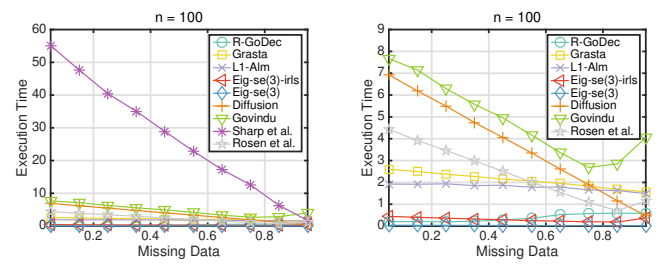


Figure 17: Execution times [seconds] as a function of $(1-p)$, with $n = 100$ and $q = 0.2$. A fixed level of noise is applied to the inlier relative motions in this experiment. The right figure is a magnification of the left one.

no outliers. In the presence of outliers, the errors achieved by such techniques slightly increase as the fraction of missing data becomes higher. As for our approach, Grasta and L1-Alm can tolerate up to 90% of missing pairs in the $q = 0.2$ case, whereas R-GoDec breaks down with 80% of missing pairs. If there are no outliers ($q = 0$), all the LRS methods can tolerate an extra 5% of missing data.

Execution Time

In this experiment we assess the computational efficiency of all the methods in two scenarios. First, we kept the density level of the measurement graph fixed ($p = 0.5$) and let n vary between 30 and 600. Then, we kept the number of absolute motions fixed ($n = 100$) and let $(1-p)$ vary between 0.05 (about 5% of missing data) and 0.95 (about 95% of missing data). In both cases we introduced a fixed level of noise and outliers on relative motions ($\sigma_T = 0.05, \sigma_R = 5^\circ$, $q = 0.2$). Diffusion is implemented in C++ (by the authors), while the remaining algorithms are implemented in Matlab. Figure 16 shows that the method by Sharp et al. (2002) is remarkably slower than the other techniques. R-GoDec, Grasta and L1-Alm are faster than Diffusion and the methods by Govindu (2004) and Rosen et al. (2016) and slower than the spectral relaxation. The R-GoDec algorithm turns out to be the fastest solution among the LRS methods, computing a solution in 12 seconds for $n = 600$. Figure 17 shows that the execution time of matrix decomposition techniques and the spectral relaxation do not change significantly when p varies, whereas the other techniques require more time as the measurement graph gets denser.

8.2. *Multiple 3D point-set registration*

In this section we report the outcome of tests on real datasets of range images. Relative motion estimates were produced thanks to the Matlab implementation of ICP (`pcregrigid`). The ICP algorithm has two outputs: an element of SE(3) that transforms one point set to the other; and a registration error computed after applying the transformation. The measurement graph was defined by discarding all the pairs with registration error higher than a threshold. This produced a redundant set of relative motions which were compensated by solving a rigid-motion synchronization problem, returning the transformations that align the original point sets. These estimates could have been improved by alternating rigid-motion synchronization and computing relative motions, as suggested by Torsello et al. (2011) and Govindu and Pooja (2014). However, such a refinement was not applied in these experiments, *i.e.* we performed rigid-motion synchronization only once.

Experimentally we observed that LRS methods perform better when translation components have values comparable to rotations, namely in the range $[-1, 1]$. For this reason, before performing rigid-motion synchronization, we divided all the relative translations by the maximum of the translations norm (and eventually multiplied the absolute translations by such a scale). This normalization also improves the results of the other algorithms.

Benchmark datasets

In this experiment we considered two benchmark repositories, which provide ground-truth absolute motions in addition to the range images. From the Stanford 3D Scanning Repository[2] we used the `Bunny`, `Happy Buddha` (standing) and `Dragon` (standing) datasets, which contain 10, 15 and 15 point sets, respectively. From the Aim@Shape-Visionair Shape Repository[3] we used the `Sheep`, `Kitten` and `Frog` datasets, which contain 20, 24 and 24 point sets, respectively. As for the initialization of the ICP algorithm, we perturbed the available ground-truth motions by a rotation with random axis and angle uniformly distributed over $[0, 2^\circ]$, similarly to the experiments carried out by Govindu and Pooja (2014).

Since ground-truth motions are available for these datasets, we evaluated quantitatively the results by reporting the mean errors in Table 3. The errors obtained by R-GoDec, Grasta, L1-Alm and EIG-SE(3)-IRLS are always lower than the other techniques, highlighting the benefit of robustness, and the worst errors are those by Sharp *et al.* We also evaluated qualitatively

[2] `http://graphics.stanford.edu/data/3Dscanrep/`
[3] `http://visionair.ge.imati.cnr.it/ontologies/shapes/`

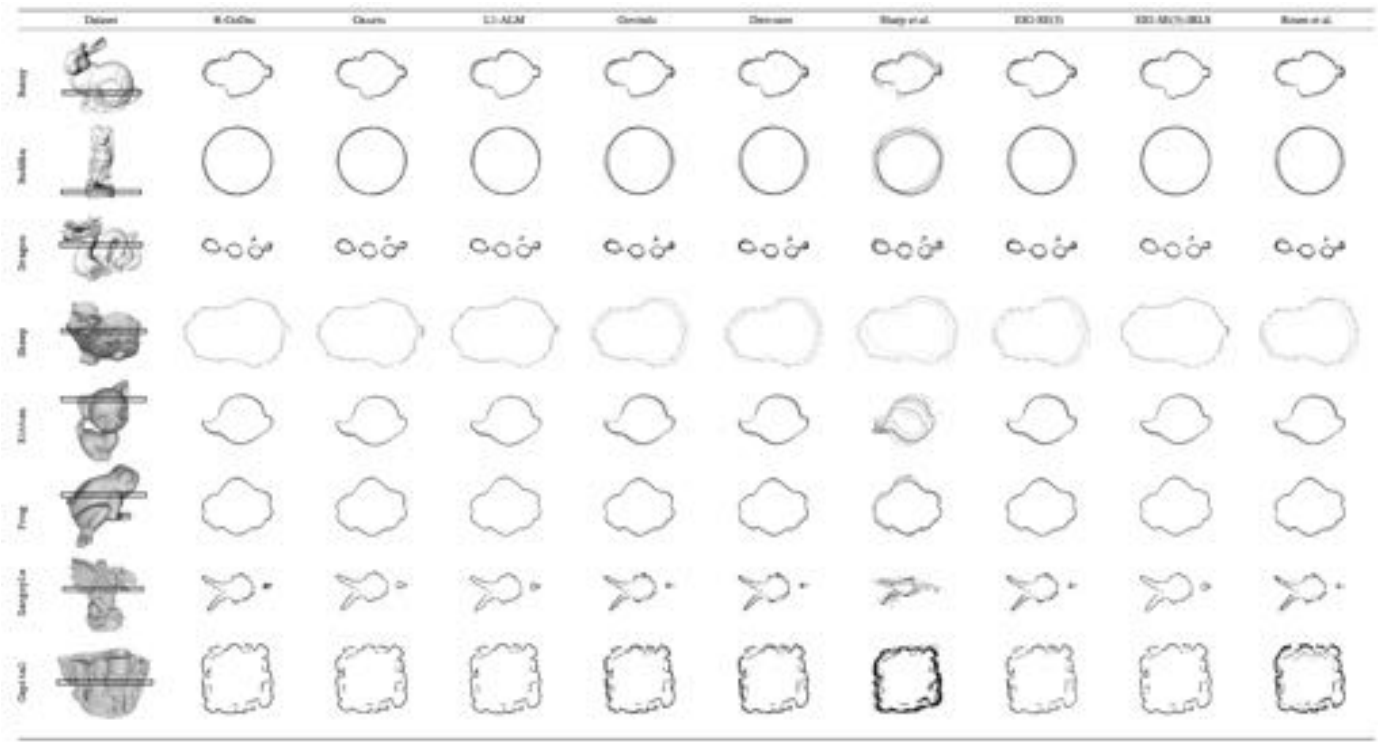
Table 3: Mean errors (rotations in degrees, translations in millimetres) on absolute motions for the Stanford and AIM@SHAPE-VISIONAIR repositories. The number of point sets and the percentage of missing pairs are also reported. The lowest errors are highlighted in boldface.

			R-GoDec		Grasta		L1-ALM		Govindu		Diffusion		Sharp *et al.*		EIG-SE(3)		EIG-SE(3)-IRLS		Rosen et al.	
Dataset	n	% miss.	rot.	tra.	rot.	tra.	rot.	tra.	rot.	tra.	rot.	tra.	rot.	tra.	rot.	tra.	rot.	tra.	rot.	tra.
Bunny	10	0	0.82	2.9	0.84	1.9	0.78	1.6	1.07	3.7	1.07	3.7	1.07	4.5	1.07	3.7	**0.73**	**1.5**	1.07	3.7
Buddha	15	0	0.85	**0.3**	0.79	0.4	0.94	0.4	1.28	0.4	1.28	0.4	2.22	0.6	1.27	0.4	**0.78**	**0.3**	1.27	0.4
Dragon	15	0	0.79	**0.3**	0.91	0.4	0.77	0.4	1.52	0.4	1.52	0.4	1.45	0.6	1.51	0.4	**0.76**	**0.3**	1.52	0.4
Sheep	20	22	0.75	0.8	0.59	1.1	0.34	0.8	5.58	4.2	5.36	4.1	3.31	4.6	4.87	3.8	**0.33**	**0.6**	5.04	3.2
Kitten	24	17	1.01	0.8	**0.90**	**0.4**	0.91	**0.4**	1.88	1.4	1.97	1.4	5.71	2.6	1.84	1.4	**0.90**	0.6	1.85	1.4
Frog	24	23	0.44	1.1	**0.26**	0.5	**0.26**	**0.4**	0.92	1.3	0.92	1.3	2.56	1.7	0.90	1.3	0.28	0.5	0.90	1.3

Table 4: Execution times (seconds) of rigid-motion synchronization. The number of point sets and the percentage of missing pairs are also reported.

Dataset	n	% miss.	R-GoDec	Grasta	L1-ALM	Govindu	Diffusion	Sharp *et al.*	EIG-SE(3)	EIG-SE(3)-IRLS	Rosen et al.
Bunny	10	0	0.02	0.17	0.08	0.04	0.09	0.30	0.04	0.34	0.15
Buddha	15	0	0.03	0.24	0.08	0.09	0.20	1.02	0.02	0.12	0.16
Dragon	15	0	0.03	0.26	0.06	0.08	0.18	0.94	0.02	0.12	0.20
Sheep	20	22	0.04	0.32	0.08	0.18	0.29	1.21	0.03	0.11	0.29
Kitten	24	17	0.03	0.44	0.11	0.20	0.42	2.15	0.05	0.13	0.30
Frog	24	23	0.04	0.41	0.10	0.15	0.39	2.44	0.05	0.17	0.27
Gargoyle	27	40	0.05	0.47	0.15	0.13	0.35	1.94	0.03	0.15	0.24
Capital	100	71	0.67	1.62	1.66	0.98	2.53	25.27	0.09	0.38	0.93

Table 5: Cross-sections of registered point-sets.



the results in terms of cross-sections in Table 5, as it is customary in the registration literature. LRS methods produce accurate cross-sections, which are comparable to EIG-SE(3)-IRLS and better than the remaining methods. The misalignment produced by Sharp et al. (2002) is particularly evident in the Kitten dataset. For visualization purposes, we report in Figure 18 the 3D models obtained by aligning the original point clouds with L1-ALM.

Execution times are reported in Table 4, which refer to the rigid-motion synchronization step, *i.e.* computing absolute motions from relative motions, and they do not include the time for computing relative motions, which is the same for all the techniques. Table 4 shows that the method by Sharp et al. (2002) is the slowest one. As for the other techniques, differences in exe-

cution times are meaningless for such relatively small datasets.

Large-scale datasets

In another experiment we considered two datasets, named `Gargoyle` and `Capital`, which contain 27 and 100 point sets respectively. Since there is no information about the scans, we simply initialized the ICP algorithm with identity matrices. This arbitrary initialization may produce some outliers among relative motions.

Execution times are reported in Table 4. R-GoDec is slower than EIG-SE(3) but faster than the other solutions, the method by Sharp et al. (2002) is the slowest technique, while Grasta and L1-Alm are faster than Diffusion, comparable to Rosen et al. (2016) and slower than EIG-SE(3)-IRLS and Govindu (2004). The different registration techniques can be appraised qualitatively from the cross-sections of output 3D models reported Table 5. The cross-sections obtained by our approach are crisper than non-robust methods, proving the effectiveness of LRS decomposition in handling measurement errors in the context of multiple 3D point-set registration. In particular, the best visual accuracy is achieved by L1-Alm and Grasta, which is comparable to EIG-SE(3)-IRLS, while R-GoDec get slightly worse results, yet better than the remaining methods. There is no significant difference between the cross-sections obtained by Diffusion, EIG-SE(3), Govindu (2004) and Rosen et al. (2016), while the misalignment produced by Sharp et al. (2002) is evident, especially for the Gargoyle dataset. Figure 18 shows the 3D models produced by L1-Alm with different colours for each point cloud.

The summary of these tests echoes those of Section 7: rigid-motion synchronization methods based on LRS decomposition provide a good trade-off between resilience to outliers and speed. However, they are more affected than the other methods by the sparsity of the graph.

9. Conclusions

After reviewing the state-of-the-art on LRS matrix decomposition, for the first time in the literature we formulated synchronization in SO(3) and SE(3) as a LRS decomposition problem. This mathematical formulation caters for missing data, outliers and noise, and it benefits from a wealth of available decomposition algorithms that can be seamlessly used as alternatives, such as Grasta or L1-Alm. Among these, we conceived R-GoDec that, building on GoDec, solves RPCA and MC together and the block structure of the problem is tailored. Experimental results show that this approach offers a good trade-off between computational efficiency and robustness to outliers. Failure cases appear when the percentage of missing data is extremely high.

Our novel formulation opens the way to the application of matrix decomposition techniques to structure-from-motion and multiple 3D point-set registration, since – in principle – any LRS algorithm can be plugged-in. In this respect, our approach will benefit from future developments in the field of LRS matrix decomposition.

Acknowledgments

The authors would like to thank Luca Magri for his support and useful suggestions. They also thank Lanhui Wang and Amit Singer for providing the Matlab code of LUD (Wang and Singer, 2013). Thanks to Massimiliano Corsini (ISTI-CNR) for providing the `Gargoyle` and `Capital` datasets, and to Avishek Chatterjee and Venu Madhav Govindu for providing the Matlab implementation of (Govindu, 2004).

References

Aftab, K., Hartley, R., Trumpf, J., 2015. Generalized weiszfeld algorithms for lq optimization. IEEE Transactions on Pattern Analysis and Machine Intelligence 4, 728 – 745.

Agrawal, M., 2006. A lie-algebraic approach for consistent pose registration for general euclidean motion, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1891–1897.

Arie-Nachimson, M., Kovalsky, S.Z., Kemelmacher-Shlizerman, I., Singer, A., Basri, R., 2012. Global motion estimation from point matches. Proceedings of the Joint 3DIM/3DPVT Conference: 3D Imaging, Modeling, Processing, Visualization and Transmission .

Arrigoni, F., Fusiello, A., Rossi, B., 2015. On computing the translations norm in the epipolar graph, in: Proceedings of the International Conference on 3D Vision (3DV), pp. 300–308.

Arrigoni, F., Fusiello, A., Rossi, B., 2016a. Camera motion from group synchronization, in: Proceedings of the International Conference on 3D Vision (3DV), pp. 546–555.

Arrigoni, F., Fusiello, A., Rossi, B., 2016b. Global registration of 3D point sets via LRS decomposition, in: Proceedings of the 14th European Conference on Computer Vision, pp. 489–504.

Arrigoni, F., Magri, L., Rossi, B., Fragneto, P., Fusiello, A., 2014a. Robust absolute rotation estimation via low-rank and sparse matrix decomposition, in: Proceedings of the International Conference on 3D Vision (3DV), pp. 491–498.

Arrigoni, F., Rossi, B., Fusiello, A., 2016c. Spectral synchronization of multiple views in SE(3). SIAM Journal on Imaging Sciences 9, 1963 – 1990.

Arrigoni, F., Rossi, B., Malapelle, F., Fragneto, P., Fusiello, A., 2014b. Robust global motion estimation with matrix completion. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-5, 63–70.

Beck, A., Teboulle, M., 2009. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences 2, 183–202.

Belta, C., Kumar, V., 2002. Euclidean metrics for motion generation on SE(3). Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science 216, 47–60.

Benjemaa, R., Schmitt, F., 1998. A solution for the registration of multiple 3D point sets using unit quaternions, in: Proceedings of the European Conference on Computer Vision, pp. 34–50.

Bernard, F., Thunberg, J., Gemmar, P., Hertel, F., Husch, A., Goncalves, J., 2015. A solution for multi-alignment by transformation synchronisation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Bertsekas, D.P., 1982. Constrained Optimization and Lagrange Multiplier Methods. Academic Press.

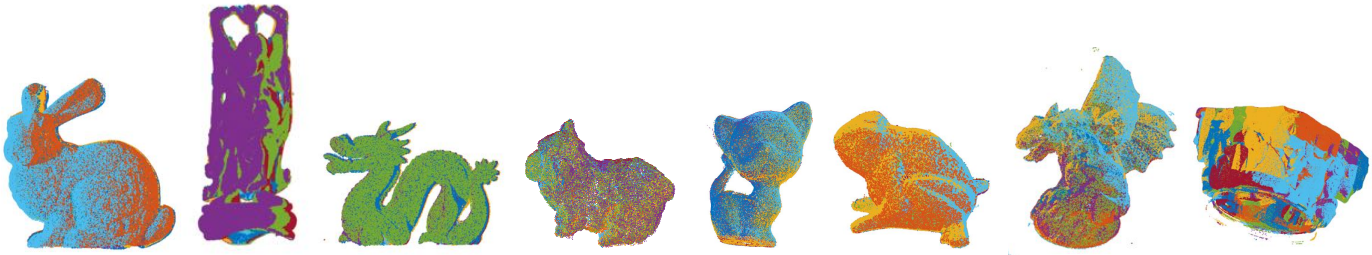
Besl, P., McKay, N., 1992. A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence 14, 239–256.

Bonarrigo, F., Signoroni, A., 2011. An enhanced 'optimization-on-a-manifold' framework for global registration of 3D range data. Proceedings of the Joint 3DIM/3DPVT Conference: 3D Imaging, Modeling, Processing, Visualization and Transmission , 350 – 357.

Boumal, N., Singer, A., Absil, P.A., 2013. Robust estimation of rotations from relative measurements by maximum likelihood, in: Proceedings of the IEEE International Conference on Robotics and Automation.

Boumal, N., Singer, A., Absil, P.A., Blondel, V.D., 2014. Cramer-Rao bounds for synchronization of rotations. Information and Inference: A Journal of the IMA 3, 1 – 39.

Bourmaud, G., 2016. Online variational Bayesian motion averaging, in: Proceedings of the European Conference on Computer Vision, pp. 126 – 142.



(a) Bunny (b) Buddha (c) Dragon (d) Sheep (e) Kitten (f) Frog (g) Gargoyle (h) Capital

Figure 18: 3D models obtained with L1-ALM. Different point sets are colour coded.

Bourmaud, G., Megret, R., Giremus, A., Berthoumieu, Y., 2014. Global motion estimation from relative measurements in the presence of outliers, in: Proceedings of the Asian Conference on Computer Vision.

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning 3, 1–122.

Brown, M., Lowe, D.G., 2005. Unsupervised 3D object recognition and reconstruction in unordered datasets, in: Proceedings of the International Conference on 3-D Digital Imaging and Modeling.

Cai, J.F., Candès, E.J., Shen, Z., 2010. A singular value thresholding algorithm for matrix completion. SIAM Journal on Optimization 20, 1956–1982.

Candès, E.J., Li, X., Ma, Y., Wright, J., 2011. Robust principal component analysis? Journal of the ACM 58, 11:1–11:37.

Candès, E.J., Recht, B., 2009. Exact matrix completion via convex optimization. Foundations of Computational Mathematics 9, 717–772.

Candès, E.J., Tao, T., 2010. The power of convex relaxation: near-optimal matrix completion. IEEE Transactions on Information Theory 56, 2053–2080.

Carlone, L., Tron, R., Daniilidis, K., Dellaert, F., 2015. Initialization techniques for 3D SLAM: A survey on rotation estimation and its use in pose graph optimization, in: Proceedings of the IEEE International Conference on Robotics and Automation.

Chatterjee, A., Govindu, V.M., 2013. Efficient and robust large-scale rotation averaging, in: Proceedings of the International Conference on Computer Vision.

Chatterjee, A., Govindu, V.M., 2017. Robust relative rotation averaging. IEEE Transactions on Pattern Analysis and Machine Intelligence .

Chaudhury, K.N., Khoo, Y., Singer, A., 2015. Global registration of multiple point clouds using semidefinite programming. SIAM Journal on Optimization 25, 468 – 501.

Chen, Y., Medioni, G., 1991. Object modeling by registration of multiple range images, in: Proceedings of the IEEE International Conference on Robotics and Automation, pp. 2724 – 2729.

Crandall, D., Owens, A., Snavely, N., Huttenlocher, D.P., 2011. Discrete-continuous optimization for large-scale structure from motion, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3001–3008.

Crosilla, F., Beinat, A., 2002. Use of generalised procrustes analysis for the photogrammetric block adjustment by independent models. ISPRS Journal of Photogrammetry & Remote Sensing 56, 195–209.

Cucuringu, M., Lipman, Y., Singer, A., 2012. Sensor network localization by eigenvector synchronization over the Euclidean group. ACM Transactions on Sensor Networks 8, 19:1 – 19:42.

Donoho, D.L., 1995. De-noising by soft-thresholding. IEEE Transactions on Information Theory 41, 613–627.

Dubbelman, G., Browning, B., 2015. COP-SLAM: Closed-form online pose-chain optimization for visual SLAM. IEEE Transactions on Robotics 31, 1194 – 1213.

Enqvist, O., Kahl, F., Olsson, C., 2011. Non-sequential structure from motion, in: Eleventh Workshop on Omnidirectional Vision, Camera Networks and Non-classical Camera.

Fantoni, S., Castellani, U., Fusiello, A., 2012. Accurate and automatic alignment of range surfaces, in: Second Joint 3DIM/3DPVT Conference: 3D Imaging, Modeling, Processing, Visualization and Transmission, pp. 73 – 80.

Fazel, M., 2002. Matrix Rank Minimization with Applications. Ph.D. thesis. Stanford University.

Fornasier, M., 2010. Theoretical Foundations and Numerical Methods for Sparse Recovery. Radon Series on Computational and Applied Mathematics, De Gruyter.

Fredriksson, J., Olsson, C., 2012. Simultaneous multiple rotation averaging using lagrangian duality, in: Proceedings of the Asian Conference on Computer Vision.

Fusiello, A., Castellani, U., Ronchetti, L., Murino, V., 2002. Model acquisition by registration of multiple acoustic range views, in: Proceedings of the European Conference on Computer Vision, pp. 805–819.

Fusiello, A., Crosilla, F., 2015. Solving bundle block adjustment by generalized anisotropic procrustes analysis. ISPRS Journal of Photogrammetry and Remote Sensing 102, 209–221.

Gherardi, R., Farenzena, M., Fusiello, A., 2010. Improving the efficiency of hierarchical structure-and-motion, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1594 – 1600.

Giridhar, A., Kumar, P., 2006. Distributed clock synchronization over wireless networks: Algorithms and analysis. Proceedings of the IEEE Conference on Decision and Control , 4915–4920.

Goldstein, T., Hand, P., Lee, C., Voroninski, V., Soatto, S., 2016. ShapeFit and ShapeKick for robust, scalable structure from motion, in: Proceedings of the European Conference on Computer Vision, pp. 289 – 304.

Govindu, V.M., 2001. Combining two-view constraints for motion estimation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Govindu, V.M., 2004. Lie-algebraic averaging for globally consistent motion estimation., in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 684–691.

Govindu, V.M., 2006. Robustness in motion averaging, in: Proceedings of the Asian Conference on Computer Vision, pp. 457–466.

Govindu, V.M., Pooja, A., 2014. On averaging multiview relations for 3D scan registration. IEEE Transactions on Image Processing 23, 1289–1302.

Hartley, R., Aftab, K., Trumpf, J., 2011. L1 rotation averaging using the Weiszfeld algorithm. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition , 3041–3048.

Hartley, R., Trumpf, J., Dai, Y., Li, H., 2013. Rotation averaging. International Journal of Computer Vision .

He, J., Balzano, L., Szlam, A., 2012. Incremental gradient on the Grassmannian for online foreground and background separation in subsampled video, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1568 – 1575.

Holland, P.W., Welsch, R.E., 1977. Robust regression using iteratively reweighted least-squares. Communications in Statistics - Theory and Methods 6, 813–827.

Jiang, N., Cui, Z., Tan, P., 2013. A global linear method for camera pose registration, in: Proceedings of the International Conference on Computer Vision.

Keller, J., 1975. Closest unitary, orthogonal and Hermitian operators to a given operator. Mathematics Magazine 48, 192–197.

Keshavan, R.H., Montanari, A., Oh, S., 2010. Matrix completion from a few entries. IEEE Transactions on Information Theory 56, 2980–2998.

Kowalski, M., Torrésani, B., 2009. Structured sparsity: from mixed norms to structured shrinkage. Signal Processing with Adaptive Sparse Structured Representations .

Krishnan, S., Lee, P.Y., Moore, J.B., Venkatasubramanian, S., 2007. Optimisation-on-a-manifold for global registration of multiple 3D point sets. International Journal of Intelligent Systems Technologies and Applications 3, 319–340.

Lee, K., Bresler, Y., 2010. ADMiRA: Atomic decomposition for minimum rank approximation. IEEE Transactions on Information Theory 56, 4402 – 4416.

Lewis, A.S., Malick, J., 2008. Alternating projections on manifolds. Mathematics of Operations Research 33, 216–234.

Lin, Z., Chen, M., Ma, Y., 2010. The augmented lagrange multiplier method for exact recovery of corrupted Low-Rank matrices. eprint arXiv:1009.5055.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60, 91–110.

Ma, S., Goldfarb, D., Chen, L., 2011. Fixed point and Bregman iterative methods for matrix rank minimization. Mathematical Programming 128, 321–353.

Martinec, D., Pajdla, T., 2007. Robust rotation and translation estimation in multiview reconstruction, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Moulon, P., Monasse, P., Marlet, R., 2013. Global fusion of relative motions for robust, accurate and scalable structure from motion, in: Proceedings of the International Conference on Computer Vision, pp. 3248–3255.

Ni, K., Dellaert, F., 2012. Hypersfm, in: Proceedings of the Joint 3DIM/3DPVT Conference: 3D Imaging, Modeling, Processing, Visualization and Transmission, pp. 144–151.

Olsson, C., Enqvist, O., 2011. Stable structure from motion for unordered image collections, in: Proceedings of the 17th Scandinavian conference on Image analysis (SCIA'11), Springer-Verlag. pp. 524–535.

Ozyesil, O., Singer, A., 2015. Robust camera location estimation by convex programming, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2674 – 2683.

Ozyesil, O., Voroninski, V., Basri, R., Singer, A., 2017. A survey of structure from motion. Acta Numerica 26, 305 – 364.

Pennec, X., 1996. Multiple registration and mean rigid shape: Applications to the 3D case, in: 16th Leeds Annual Statistical Workshop, pp. 178–185.

Peters, J.R., Borra, D., Paden, B.E., Bullo, F., 2015. Sensor network localization on the group of three-dimensional displacements. SIAM Journal on Control and Optimization 53, 3534–3561.

Piovan, G., Shames, I., Fidan, B., Bullo, F., Anderson, B., 2013. On frame and orientation localization for relative sensing networks. Automatica 49, 206–213.

Pulli, K., 1999. Multiview registration for large data sets, in: Proceedings of the International Conference on 3-D Digital Imaging and Modeling, pp. 160–168.

Rosen, D., Carlone, L., Bandeira, A., Leonard, J., 2016. A certifiably correct algorithm for synchronization over the special Euclidean group, in: International Workshop on the Algorithmic Foundations of Robotics (WAFR).

Rosen, D.M., DuHadway, C., Leonard, J.J., 2015. A convex relaxation for approximate global optimization in simultaneous localization and mapping, in: Proceedings of the IEEE International Conference on Robotics and Automation, pp. 5822 – 5829.

Rossi, B., Patanè, M., Fragneto, P., Fusiello, A., 2017. Robust localization in wireless sensor networks via low-rank and sparse matrix decomposition. International Conference on Future Networks and Distributed Systems .

Rusinkiewicz, S., Levoy, M., 2001. Efficient variants of the ICP algorithm, in: Proceedings of the International Conference on 3-D Digital Imaging and Modeling, pp. 145–152.

Saunderson, J., Parrilo, P.A., Willsky, A.S., 2015. Semidefinite descriptions of the convex hull of rotation matrices. SIAM Journal on Optimization 25, 1314 – 1343.

Sharp, G.C., Lee, S.W., Wehe, D.K., 2002. Multiview registration of 3D scenes by minimizing error between coordinate frames, in: Proceedings of the European Conference on Computer Vision, pp. 587–597.

Shih, S., Chuang, Y., Yu, T., 2008. An efficient and accurate method for the relaxation of multiview registration error. IEEE Transactions on Image Processing 17, 968 – 981.

Singer, A., 2011. Angular synchronization by eigenvectors and semidefinite programming. Applied and Computational Harmonic Analysis 30, 20 – 36.

Singer, A., Shkolnisky, Y., 2011. Three-dimensional structure determination from common lines in cryo-em by eigenvectors and semidefinite programming. SIAM Journal on Imaging Sciences 4, 543 – 572.

Snavely, N., Seitz, S.M., Szeliski, R., 2006. Photo tourism: exploring photo collections in 3d, in: SIGGRAPH: International Conference on Computer Graphics and Interactive Techniques, pp. 835–846.

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, pp. 1–8.

Sturm, J.F., 1999. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Optimization Methods and Software 11-12, 525–653.

Tao, M., Yuan, X., 2011. Recovering low-rank and sparse components of matrices from incomplete and noisy observations. SIAM Journal on Optimization 21, 57–81.

Toldo, R., Beinat, A., Crosilla, F., 2010. Global registration of multiple point clouds embedding the generalized procrustes analysis into an ICP framework, in: International Symposium on 3D Data Processing, Visualization and Transmission.

Torsello, A., Rodolá, E., Albarelli, A., 2011. Multiview registration via graph diffusion of dual quaternions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2441 – 2448.

Triggs, B., McLauchlan, P.F., Hartley, R.I., Fitzgibbon, A.W., 2000. Bundle adjustment - a modern synthesis, in: Proceedings of the International Workshop on Vision Algorithms, Springer-Verlag. pp. 298–372.

Tron, R., Danilidis, K., 2014. Statistical pose averaging with varying and non-isotropic covariances, in: Proceedings of the European Conference on Computer Vision.

Tron, R., Vidal, R., 2014. Distributed 3-D localization of camera sensor networks from 2-D image measurements. IEEE Transactions on Automatic Control 59, 3325–3340.

Tron, R., Zhou, X., Daniilidis, K., 2016. A survey on rotation optimization in structure from motion, in: Computer Vision and Pattern Recognition Workshops (CVPRW).

Tropp, A., Halko, N., Martinsson, P., 2010. Finding structure with randomness: Stochastic algorithms for constructing approximate matrix decompositions.

Tseng, P., 2001. Convergence of a block coordinate descent method for non-differentiable minimization. Journal of optimization theory and applications 109, 475–494.

Wang, L., Singer, A., 2013. Exact and stable recovery of rotations for robust synchronization. Information and Inference: a Journal of the IMA 2, 145–193.

Wang, Y.X., Lee, C.M., Cheong, L.F., Toh, K.C., 2014. Practical matrix completion and corruption recovery using proximal alternating robust subspace minimization. International Journal of Computer Vision , 1–30.

Waters, A.E., Sankaranarayanan, A.C., Baraniuk, R.G., 2011. Sparcs: Recovering low-rank and sparse matrices from compressive measurements, in: Neural Information Processing Systems.

Wen, Z., Yin, W., Zhang, Y., 2012. Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm. Mathematical Programming Computation 4, 333–361.

Wilson, K., Bindel, D., Snavely, N., 2016. When is rotations averaging hard?, in: Proceedings of the European Conference on Computer Vision, pp. 255 – 270.

Wilson, K., Snavely, N., 2014. Robust global translations with 1DSfM, in: Proceedings of the European Conference on Computer Vision, pp. 61–75.

Zach, C., Klopschitz, M., Pollefeys, M., 2010. Disambiguating visual relations using loop constraints, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1426 – 1433.

Zhao, S., Zelazo, D., 2016. Localizability and distributed protocols for bearing-based network localization in arbitrary dimensions. Automatica 69, 334 – 341.

Zheng, Y., Liu, G., Sugimoto, S., Yan, S., Okutomi, M., 2012. Practical low-rank matrix approximation under robust L_1-norm, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1410–1417.

Zhou, T., Tao, D., 2011. GoDec: Randomized low-rank & sparse matrix decomposition in noisy case, in: Proceedings of the 28th International Conference on Machine Learning (ICML11), ACM. pp. 33–40.

Zhou, T., Tao, D., 2013. Shifted subspaces tracking on sparse outlier for motion segmentation., in: IJCAI, pp. 1946–1952.

Zhou, X., Yang, C., Zhao, H., Yu, W., 2014. Low-rank modeling and its applications in image analysis. ACM Computing Surveys 47, 36:1–36:33.

Zhou, Z., Li, X., Wright, J., Candès, E.J., Ma, Y., 2010. Stable principal component pursuit, in: IEEE International Symposium on Information Theory, pp. 1518–1522.