# Global Motion Estimation from Relative Measurements in the Presence of Outliers

Guillaume Bourmaud[1,2], Rémi Mégret[1,2,3],
Audrey Giremus[1,2], Yannick Berthoumieu[1,2]

[1] Univ. Bordeaux, IMS, UMR 5218, F-33400 Talence, France.
[2] CNRS, IMS, UMR 5218, F-33400 Talence, France.
[3] Dept of Mathematical Sciences, University of Puerto-Rico at Mayaguez

**Abstract.** This work addresses the generic problem of global motion estimation (homographies, camera poses, orientations, etc.) from relative measurements in the presence of outliers. We propose an efficient and robust framework to tackle this problem when motion parameters belong to a Lie group manifold. It exploits the graph structure of the problem as well as the geometry of the manifold. It is based on the recently proposed iterated extended Kalman filter on matrix Lie groups. Our algorithm iteratively samples a minimum spanning tree of the graph, applies Kalman filtering along this spanning tree and updates the graph structure, until convergence. The graph structure update is based on computing loop errors in the graph and applying a proposed statistical inlier test on Lie groups. This is done efficiently, taking advantage of the covariance matrix of the estimation errors produced by the filter. The proposed formalism is applied on both synthetic and real data, for a camera pose registration problem, an automatic image mosaicking problem and a partial 3D reconstruction merging problem. In these applications, the framework presented in this paper efficiently recovers the global motions while the state of the art algorithms fail due to the presence of a large proportion of outliers.

## 1 Introduction

This paper deals with the generic problem of estimating globally consistent motion parameters (global motions) from relative motion measurements in the presence of outliers. Such a problem occurs for instance in the context of camera pose registration [1] encountered in 3D localization, structure from motion, camera network calibration, etc. In this case, a motion or transformation is a rigid body transformation matrix. Thus, the relative measurements correspond to the rigid transformations between two cameras and the global motions we wish to estimate are the rigid transformation matrices between a reference camera and all the other cameras. For this specific application, two different kinds of outlier measurements can occur. The first kind of outliers are statistically independent from each other and arise from random failures such as RANSAC [2] failure, erroneous matches between pair of images, etc. The second kind of outliers are not independent and are due to duplicated structures in the environment [3]. For example, 3 images taken in 3 different places that are very similar match each other and thus produce 3 outlier relative motion measurements that are coherent with each other.

The generic problem considered in this work has several other applications such as multiple rotation averaging [4] (3-dimensional rotation matrices), image mosaicking [5] (3-dimensional homographies) and partial 3D reconstruction merging [6] (4-dimensional similarity transformation matrices).

All these applications can be seen as an inference problem in a pairwise factor graph (PFG) where both the vertices, i.e the global motions, and the edges, i.e the noisy relative motions, evolve on a matrix Lie group [7]. Indeed, a 3D rotation matrix evolves on the Lie group $SO(3)$ [7], the rigid body motion matrices correspond to the Special Euclidean Lie group $SE(3)$ [8], the homography matrices can be identified with the Special Linear Lie group $SL(3)$ [9], and the 3D similarity transformation matrices form the Lie group $Sim(3)$ [10].

In these applications, as explained above for the camera pose registration problem, the outlier measurements might not be independent and frequently outnumber the inlier measurements. As a consequence, robust optimization approaches, such as Huber norm [11], that do not explicitly exclude the outliers from the estimation process, typically fail. Without additional information, such as priors on whether a relative motion measurement is inlier or outlier, the solution forming the largest coherent set of relative motions may include dependent outliers. Consequently, the global motions are not correctly recovered.

The formalism proposed in this paper can deal with any matrix Lie group and includes the a priori information on whether a relative motion measurement is inlier or outlier as a weighted adjacency matrix (WAM) of the PFG. Consequently, it is able to tackle each of the previously mentioned applications while excluding the outliers in the relative motion measurements. It relies on the recently proposed Iterated Extended Kalman Filter on Lie Groups (LG-IEKF) [12]. Our algorithm iteratively samples a minimum spanning tree (MST) in the WAM of the PFG, applies the LG-IEKF along this MST and updates the WAM, until convergence. The WAM update is based on computing loop errors in the graph and applying a proposed statistical inlier test on Lie groups. This is performed efficiently, taking advantage of the covariance matrix of the estimation errors produced by the LG-IEKF.

The rest of the paper is organized as follows: the next section deals with related work. Section 3 introduces the formalism of Lie groups. The proposed framework is described in section 4. In section 5, our formalism is evaluated experimentally on several applications. Finally the conclusion is provided in section 6.

## 2 Related Work

A large amount of work has been recently devoted to specifically dealing with multiple rotation averaging in the presence of outliers. This problem is also known as synchronization of rotations in the mathematics community and is usually tackled by minimizing a given criterion. In [13] and [14], spectral relaxations of the problem are proposed while [15] uses their results as initialization for a second order Riemannian trust-region algorithm to compute a local maximizer. [16] derives an algorithm that exactly estimates the global rotations when a subset of the measurements are perfect and outperforms [14]. In [17] and [18], two robust iterative algorithms, based on L1 and L1-L2 minimization criterion, respectively, are devised. However, the considered error functions are not convex and consequently need a good initialization such as [16] to avoid poor local minima. Finally, [19] proposes a discretization of $SO(3)$ to apply a loopy belief propagation algorithm on the resulting Markov random field.

All the previously cited approaches, assume that the outliers are statistically independent. Consequently, none of them is able to correctly recover the global motions when this assumption is violated. The works [20], [21] and [22] are also relevant for the multiple rotation averaging problem. However, they are specifically tailored for $SO(3)$ and it is not straightforward to apply them to other Lie groups.

In [23], a method, that also assumes independent outliers, is derived to infer the set of outliers. The authors introduce a Bayesian framework based on collecting the loop errors in the PFG to infer outliers. Unfortunately, collecting the loop errors becomes quickly intractable and the maximum loop length is limited to 6. Consequently, many outliers cannot be detected (see [20] Fig.4). Limiting the maximum loop length to 6 also allows them not to take into account the uncertainty induced by the length of a loop. Moreover, when the outliers are not independent, it is possible to find loops containing outliers that have a very low loop error. Thus, in this case, the method fails to infer the dependent outliers as it was shown on several examples in [24] and [3].

To the best of our knowledge, only one approach [3] was proposed to deal with the generic problem of global motion estimation from relative measurements in the presence of statistically dependent outliers. It is inspired by [25] which proposed a RANSAC-like algorithm to estimate the global motions. It consists in drawing spanning trees (ST) in the PFG. However, using random sampling, the number of ST to draw before finding an outlier free ST is huge. Thus, [3] proposes to sample the STs from a WAM in order to increase the chances to draw an ST without outliers. For each sampled ST, an Expectation Maximization (EM) algorithm is applied, introducing latent variables to classify the measurements as inliers or outliers. Finally, from these labels, a likelihood based on the weights of the WAM is defined and the solution of the ST which maximizes this likelihood is chosen. This approach is shown to perform very well on several small datasets. However, as we show in our experiments, it can only be applied when the number of global motions is small. Moreover, the proposed EM algorithm, although initialized with an outlier free ST, can converge to a poor local minimum (see section 5).

[26] is also a relevant work dedicated to large scale problems though, as specified by the authors themselves, this method "cannot disambiguate" as well as [3]. Consequently, in the rest of the paper, this approach is not considered.

In robotics, several relevant works [27–29] dedicated to robust graph SLAM, have been published. However, they assume that an outlier free ST is given. Furthermore, neither parameter estimation on Lie group nor measurements on Lie group is addressed.

In this paper, we propose a generic approach combining a sampling approach, as in [25], the use of a WAM, as in [3], and the computation of loop errors, as in [23]. Based on these three ingredients, the recently proposed LG-IEKF [12] and a proposed statistical inlier test on Lie groups, we derive an efficient algorithm which is able to recover the global motions in the presence of statistically dependent outliers when the state of the art algorithms, previously cited, fail.

## 3 Preliminaries

*Introduction to matrix Lie groups* In this section, we briefly introduce the matrix Lie Groups for the specific purpose of transformation/motion estimation. For a detailed description of these notions the reader is referred to [7]. If $G$ is a matrix Lie group, then $X_{ij} \in G \subset \mathbb{R}^{n \times n}$ is a transformation matrix that takes a point $x^j \in \mathbb{R}^n$ defined in the reference frame (RF) $j$ to RF $i$, i.e $x^i = X_{ij} x^j$. Two transformations $X_{ij} \in G$ and $X_{jk} \in G$ can be composed using matrix multiplication to obtain another transformation $X_{ik} = X_{ij} X_{jk} \in G$. Inverting a transformation matrix $X_{ij}$ produces the inverse transformation, i.e $X_{ij}^{-1} = X_{ji}$. Consequently multiplying a transformation with its inverse produces the identity matrix: $X_{ij} X_{ji} = Id_{n \times n}$. The matrix exponential $exp_G$ and matrix logarithm $log_G$ mappings establish a local diffeomorphism between an open neighborhood of $\mathbf{0}_{n \times n}$ in the tangent space at the identity $T_e G$, called the *Lie Algebra* $\mathfrak{g}$, and an open neighborhood of $Id_{n \times n}$ in $G$. The Lie Algebra $\mathfrak{g}$ associated to a $p$-dimensional matrix Lie group is a $p$-dimensional vector space. Hence there is a linear isomorphism between $\mathfrak{g}$ and $\mathbb{R}^p$ that we denote as follows: $[\cdot]_G^\vee : \mathfrak{g} \to \mathbb{R}^p$ and $[\cdot]_G^\wedge : \mathbb{R}^p \to \mathfrak{g}$. We also introduce the following notations: $exp_G^\wedge(\cdot) = exp_G\left([\cdot]_G^\wedge\right)$ and $log_G^\vee(\cdot) = [log_G(\cdot)]_G^\vee$. It means that a transformation $X_{jj'}$ that is "close enough" to $Id_{n \times n}$ can be parametrized as follows: $X_{jj'} = exp_G^\wedge(\delta_{jj'}) \in \mathbb{R}^p$. Finally, we remind the adjoint representation $Ad_G(\cdot) \subset \mathbb{R}^{p \times p}$ of $G$ on $\mathbb{R}^p$ that enables us to transport an increment $\epsilon_{ij}^i \in \mathbb{R}^p$, that acts onto an element $X_{ij}$ through left multiplication, into an increment $\epsilon_{ij}^j \in \mathbb{R}^p$, that acts through right multiplication:

$$exp_G^\wedge\left(\epsilon_{ij}^i\right) X_{ij} = X_{ij} exp_G^\wedge\left(Ad_G\left(X_{ij}^{-1}\right)\epsilon_{ij}^i\right) = X_{ij} exp_G^\wedge\left(\epsilon_{ij}^j\right) \tag{1}$$

where $\epsilon_{ij}^j = Ad_G\left(X_{ij}^{-1}\right)\epsilon_{ij}^i = Ad_G\left(X_{ji}\right)\epsilon_{ij}^i$.

*Concentrated Gaussian Distribution on Lie Groups* In this section, we briefly introduce the concept of concentrated Gaussian on Lie groups [30–33] as a generalization of the normal distribution to Lie group manifolds. The distribution of

$X_{ij} \in G$ is called a (right) concentrated Gaussian distribution on $G$ of "mean" $\mu_{ij}$ and "covariance" $P_{ii}$ denoted $X_{ij} \sim \mathcal{N}_G^R(\mu_{ij}, P_{ii})$ if:

$$X_{ij} = exp_G^{\wedge}\left(\epsilon_{ij}^i\right)\mu_{ij} = \mu_{ij}exp_G^{\wedge}\left(Ad_G\left(\mu_{ij}^{-1}\right)\epsilon_{ij}^i\right) = \mu_{ij}exp_G^{\wedge}\left(\epsilon_{ij}^j\right) \tag{2}$$

where $\epsilon_{ij}^i \sim \mathcal{N}_{\mathbb{R}^p}\left(\mathbf{0}_{p\times 1}, P_{ii}\right)$, $\epsilon_{ij}^j \sim \mathcal{N}_{\mathbb{R}^p}\left(\mathbf{0}_{p\times 1}, Ad_G\left(\mu_{ij}^{-1}\right)P_{ii}Ad_G\left(\mu_{ij}^{-1}\right)^T\right)$ and $P_{ii} \subset \mathbb{R}^{p\times p}$ is a definite positive matrix. Such a distribution gives us a meaningful covariance representation. In the rest of the paper, it will allow us to quantify the uncertainty of both the global and the relative motions and thus to statistically define a threshold to reject outlier measurements.

## 4 Global Motion Estimation from Relative Measurements in the Presence of Outliers

In this work, we aim at estimating global motions $\{X_{iR}\}_{i=1:N}$, where each global motion $X_{iR} \in G^{'}$ is defined as the motion between a main RF $R$ and a RF $i$, and $G^{'}$ is a $p$-dimensional matrix Lie group such as $SO(3)$, $SE(3)$, $SL(3)$, $Sim(3)$, etc. An illustration of global motions in the context of an outlier free consistent pose registration problem (Lie group $SE(3)$) is presented in Fig.1a. First of all, we describe the case where relative motion measurements are not corrupted with outliers. Then, we treat the case of robust estimation.

### 4.1 Outlier Free Estimation

This section is mainly a summary of [12], however, its understanding is mandatory for the rest of the paper.

**Model** We consider the case where the noises on the (inlier) relative motion measurements $\{Z_{ij}\}_{1\leq i<j\leq N}$ are mutually independent. Each $Z_{ij} \in G^{'}$ denotes a noisy relative motion between a RF $j$ and a RF $i$ expressed as follows:

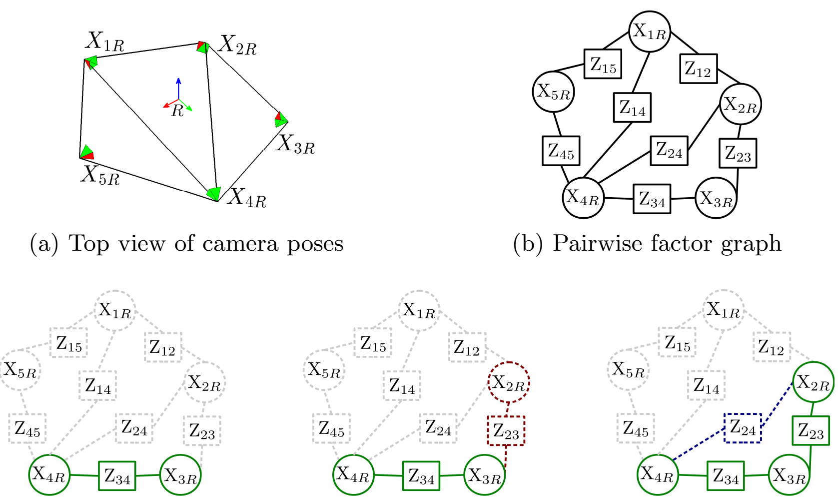
$$Z_{ij} = exp_G^{\wedge}\left(b_{ij}^i\right)X_{iR}X_{jR}^{-1} \tag{3}$$

where $b_{ij}^i \sim \mathcal{N}_{\mathbb{R}^p}\left(\mathbf{0}_{p\times 1}, \Sigma_{ii}\right)$ is a white Gaussian noise. The problem considered can be seen as the inference in a PFG $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where each vertex $\mathcal{V}_i$ corresponds to a global motion $X_{iR}$ and each pairwise factor $\mathcal{E}_{ij}$ corresponds to a relative measurement $Z_{ij}$ (see Fig.1b). In this paper, $\mathcal{G}$ denotes either the PFG itself or its (weighted) adjacency matrix.

Under the concentrated Gaussian assumption, the maximum likelihood estimates of the global motions denoted $\{\mu_{iR}\}_{i=1:N}$ are then defined as:

$$\{\mu_{iR}\}_{i=1:N} = \underset{\{X_{iR}\}_{i=1:N}}{\operatorname{argmin}}\left(\sum_{i,j}\left\|log_{G'}^{\vee}\left(Z_{ij}X_{jR}X_{iR}^{-1}\right)\right\|_{\Sigma_{ii}}^2\right) \tag{4}$$

where $\|\cdot\|_{\Sigma}^2$ stands for the squared Mahalanobis distance.




(a) Top view of camera poses (b) Pairwise factor graph

(c) Initialization (d) Propagation (e) Update

Fig. 1: Illustration of an outlier free consistent pose registration problem (Lie group $SE(3)$): (a) a cone represents a camera (global motion) and a link between two cones indicates that a relative motion measurement is available. (b)-(e) please see explanations in section 4.1

**Iterated Extended Kalman Filter on Matrix Lie Groups** The problem considered in (4) has several local minima and an efficient way to reach a "good" local minimum is to apply an LG-IEKF (see [12]). The idea is to draw an ST from the adjacency matrix $\mathcal{G}$ of the PFG that guides the global motion estimates at each propagation step of the filter. At each update step of the filter, the relative motions that close loops in the graph are used to refine the global motion estimates and reduce their uncertainty.

**Spanning Tree $\mathcal{T}$:** Let's consider an ST of the PFG $\mathcal{G} : \mathcal{T} = \{\mathcal{V}, \mathcal{E}_{\mathcal{T}}\}$. $\mathcal{E}_{\mathcal{T}} = (C^m)_{m=0:N-2}$ corresponds to a tree traversal ordered such that $C^n = Z_{i_{\mathcal{T}}(n) j_{\mathcal{T}}(n)}$ is connected to the tree built from $(C^m)_{m=0:n-1}$. The index $m$ can be seen as a time instant and will be referred as such in the rest of the paper. The notations $i_{\mathcal{T}}(m)$ and $j_{\mathcal{T}}(m)$ indicate the referential frames $i$ and $j$ associated to the relative measurement $C^m$.

**Loop Closure $\mathcal{L}$:** In this context, a loop closure (LC) at time instant $n$ is a relative measurement $Z_{ij} \notin \mathcal{E}_{\mathcal{T}}$ connected to the tree built from $(C^m)_{m=0:n+1}$ and not connected to the tree built from $(C^m)_{m=0:n}$. We define the ordered LCs as $\mathcal{L} = (M^m)_{m=1:N-2}$. Note that the size of $M^m$ depends on the time instant $m$. Indeed, for instance, if two LC occur at time instant $m$, then $M^m = \left\{Z_{i_{\mathcal{L}}(m,1) j_{\mathcal{L}}(m,1)}, Z_{i_{\mathcal{L}}(m,2) j_{\mathcal{L}}(m,2)}\right\}$ is a set that contains two relative measurements. The notations $i_{\mathcal{L}}(m,\mathfrak{z})$ and $j_{\mathcal{L}}(m,\mathfrak{z})$ indicate the referential frames $i$ and $j$ for the $\mathfrak{z}^{th}$ LC of $M^m$. We introduce the following notation:

$M_{\mathfrak{z}}^{m} = Z_{i_{\mathcal{L}}(m,\mathfrak{z}) j_{\mathcal{L}}(m,\mathfrak{z})}$ which means that at time instant $m$, the $\mathfrak{z}^{th}$ LC is a transformation from RF $j_{\mathcal{L}}(m,\mathfrak{z})$ to RF $i_{\mathcal{L}}(m,\mathfrak{z})$. From an implementation point of view, the variables $i_{\mathcal{T}}$, $j_{\mathcal{T}}$, $i_{\mathcal{L}}$ and $j_{\mathcal{L}}$ are tables of indices indicating the RFs associated to the relative transformation measurements in $\mathcal{E}_{\mathcal{T}}$ and $\mathcal{L}$.

**Measurement Covariance:** In the two previous paragraphs, we have introduced notations to distinguish a relative measurement $Z_{i_{\mathcal{T}}(n) j_{\mathcal{T}}(n)}$ that is part of $\mathcal{E}_{\mathcal{T}}$ from a relative measurement $Z_{i_{\mathcal{L}}(m,\mathfrak{z}) j_{\mathcal{L}}(m,\mathfrak{z})}$ that is part of $\mathcal{L}$. All these measurements arise from the same generative model (3). Thus the covariance matrix of $Z_{i_{\mathcal{T}}(n) j_{\mathcal{T}}(n)}$ is noted $\Sigma_{i_{\mathcal{T}}(n) i_{\mathcal{T}}(n)}$ whereas the covariance matrix of $Z_{i_{\mathcal{L}}(m,\mathfrak{z}) j_{\mathcal{L}}(m,\mathfrak{z})}$ is noted $\Sigma_{i_{\mathcal{L}}(m,\mathfrak{z}) i_{\mathcal{L}}(m,\mathfrak{z})}$.

**Scheduling:** A scheduling is a choice of $\mathcal{E}_{\mathcal{T}}$ for a given graph $\mathcal{G}$. A possible scheduling for the graph presented Fig.1b is: $\mathcal{E}_{\mathcal{T}} = \left(C^0 = Z_{34}, C^1 = Z_{23}, C^2 = Z_{45}, C^3 = Z_{14}\right)$ . It implies the following set of LC: $\mathcal{L} = \left(M^1 = Z_{24}, M^2 = \{\emptyset\}, M^3 = \left\{M_1^3, M_2^3\right\} = \{Z_{12}, Z_{15}\}\right)$

**Algorithm:** Once a scheduling is decided, the LG-IEKF algorithm (see [12]) can be applied to estimate both the global motions $\mu$ and the covariance of the estimation errors $P$. Step 1 (Initialization), 2 (Propagation) and 3 (Update) of the LG-IEKF algorithm are illustrated Fig. 1c, 1d and 1e for the scheduling previously defined.

### 4.2 Estimation in the presence of outliers

The LG-IEKF algorithm described in the previous section is not robust to outliers in the relative motion measurements all the more the noise is modeled as a concentrated Gaussian distribution on Lie groups (see (3)). In this section, we show how to perform the estimation in the presence of outliers.

**Outlier Definition and Inlier Test** As previously explained, the outliers arising in the problem we consider in this paper can be statistically dependent. It is a difficult task to propose a generative model in this case. Consequently, in this work, we simply use a discriminative way to define an outlier. A relative motion measurement $Z_{ij}$ is an outlier if and only if:

$$\left\| log_{G'}^{\vee}\left(Z_{ij} X_{jR} X_{iR}^{-1}\right)\right\|_{\Sigma_{ii}}^{2} > thresh \tag{5}$$

where $thresh$ is a threshold to be defined. Note that, with this definition, a relative measurement generated using the inlier model (3) can be classified as outlier. However, if $thresh$ is large enough, it is very unlikely to happen.

From the outlier definition given in (5), we propose a statistical inlier test on matrix Lie groups that will be employed in our robust estimation framework: let's assume that we have estimated the two global motions $X_{iR}$ and $X_{jR}$, as well as the covariance of the estimation errors from the relative motion measurements $\{Z_{kl}\}_{(k,l)\in T}$ ($T$ is a subset of all the measurements) without involving the relative measurement $Z_{ij}$, i.e $(i,j) \notin T$. We would like to know whether $Z_{ij}$ is an "inlier" or not w.r.t the current estimates of $X_{iR}$ and $X_{jR}$. Assuming that the distribution of the two global motions $X_{iR}$ and $X_{jR}$ conditioned by the relative measurements $\{Z_{kl}\}_{(k,l)\in T}$ is a Gaussian distribution on Lie groups, we have:

$$X_{iR}|\{Z_{kl}\}_{(k,l)\in T} = exp^{\wedge}_{G'}\left(\epsilon^{i}_{iR}\right)\mu_{iR} \text{ and } X_{jR}|\{Z_{kl}\}_{(k,l)\in T} = exp^{\wedge}_{G'}\left(\epsilon^{j}_{jR}\right)\mu_{jR} \tag{6}$$

where $cov(\epsilon) = cov\left(\begin{bmatrix}\epsilon^{i}_{iR}\\ \epsilon^{j}_{jR}\end{bmatrix}\right) = \begin{bmatrix}P_{ii} & P_{ij}\\ P_{ji} & P_{jj}\end{bmatrix}$. From (3), we have the following result:

$$Z_{ij} = exp^{\wedge}_{G}\left(b^{i}_{ij}\right) exp^{\wedge}_{G'}\left(\epsilon^{i}_{iR}\right)\mu_{iR}\mu^{-1}_{jR} exp^{\wedge}_{G'}\left(-\epsilon^{j}_{jR}\right) \tag{7}$$

$$= exp^{\wedge}_{G}\left(b^{i}_{ij} + \epsilon^{i}_{iR} - Ad_{G'}\left(\mu_{iR}\mu^{-1}_{jR}\right)\epsilon^{j}_{jR} + O\left(\|\epsilon\|^{2}, \left\|b^{i}_{ij}\right\|^{2}\right)\right)\mu_{iR}\mu^{-1}_{jR} \tag{8}$$

Thus, neglecting second order terms, the error negative log-likelihood has the following expression:

$$err = \left\|log^{\vee}_{G'}\left(Z_{ij}\mu_{jR}\mu^{-1}_{iR}\right)\right\|^{2}_{Q_{err}} \tag{9}$$

where

$$Q_{err} = cov\left(\epsilon^{i}_{iR} - Ad_{G'}\left(\mu_{iR}\mu^{-1}_{jR}\right)\epsilon^{j}_{jR} + b^{i}_{ij}\right) \tag{10}$$

$$= \left[Id_{p\times p} \; -Ad_{G'}\left(\mu_{iR}\mu^{-1}_{jR}\right)\right]\begin{bmatrix}P_{ii} & P_{ij}\\ P_{ji} & P_{jj}\end{bmatrix}\begin{bmatrix}Id_{p\times p}\\ -Ad_{G'}\left(\mu_{iR}\mu^{-1}_{jR}\right)^{T}\end{bmatrix} + \Sigma_{ii} \tag{11}$$

and is distributed according to the chi-squared distribution with $p$ degrees of freedom, i.e $err \sim \chi^{2}(p)$. Consequently, one way to decide whether $Z_{ij}$ is an inlier w.r.t the current estimates of $X_{iR}$ and $X_{jR}$ is to define a threshold based on the p-value of $\chi^{2}(p)$ [34]. Note that since we neglected second order terms, this threshold is possibly restrictive, thus in practice we take a larger threshold than the theoretical one.

**Loop Voting** The robust framework that is presented in the next section is based on "loop voting". The idea of "loop voting" is simple: assuming an ST $\mathcal{E}_{\mathcal{T}} = (C^{m})_{m=0:N-2}$, which probably contains outliers, has been drawn from $\mathcal{G}$, we apply the LG-IEKF algorithm. At time instant $k$, between Step 2 and Step 3 of this algorithm, we perform the inlier test described in section 4.2 for the upcoming loop closures $M^{k}$. For simplicity let's consider the first loop closure $M^{k}_{i_{\mathcal{L}}(k,\imath)j_{\mathcal{L}}(k,\imath)}$, moreover we define $\mathfrak{m} = i_{\mathcal{L}}(k,\imath)$ and $\mathfrak{n} = j_{\mathcal{L}}(k,\imath)$. If the inlier test, for this loop closure, is validated, i.e $\left\|log^{\vee}_{G'}\left(Z_{\mathfrak{mn}}\mu^{k|k-1}_{\mathfrak{n}R}\left(\mu^{k|k-1}_{\mathfrak{m}R}\right)^{-1}\right)\right\|^{2}_{Q_{err}} < thresh$, then we found a measurement that is "coherent" with the path $\mathcal{P}$ in $\mathcal{E}_{\mathcal{T}}$ from RF $\mathfrak{m}$ to RF $\mathfrak{n}$. Thus we mark the relative motions of $\mathcal{P}$ as "checked"[1]. Note that, if $\mathcal{E}_{\mathcal{T}}$ contains dependent

[1] We use the Matlab library Matgraph [35] to find the path between $\mathfrak{m}$ and $\mathfrak{n}$.

outliers (from duplicate structures in the scene for instance), then it is possible to close wrong loops and thus to "check" outliers. This point is discussed in the next two sections.

**Proposed Framework** In order to perform the robust estimation of global motions from relative motions in the presence of statically dependent outliers, we proposed to combine three ideas:

- *sampling Minimum Spanning Trees (MST) from a weighted adjacency matrix (WAM), as in [25, 3].* The WAM contains prior information on whether a relative measurement is inlier or outlier. Without this prior information, the solution forming the largest coherent set of relative motions may include dependent outliers. However, even with a WAM, the combinatorial search of the "best" relative motions configuration is intractable. In [3], it is proposed to sample a large number of spanning trees from the WAM, to perform an optimization on each ST and to keep the best solution. Nevertheless, we show in section 5 that their approach can only be applied for a small number of global motions. In order to obtain an algorithm that is applicable for larger problems, we assume that an MST of $\mathcal{G}$ does not contain dependent outliers (it may contain independent outliers). This assumption might appear restrictive, however, for image sequences for example, it is usually satisfied (see section 4.2). Indeed, the image timestamps can be used to build a WAM that favors images that are close in time. In this case, an MST will not contain dependent outliers since the relative motion measurement between two consecutive images is normally either an inlier or a statistically independent outlier (in the case of a RANSAC failure for example). When timestamps are not available, one way to attribute weights to the edges is to consider "missing correspondence cue" (see [3]). Note that if there is only statistically independent outliers, our approach does not need a WAM.

- *applying the LG-IEKF proposed in [12].* This algorithm achieves similar performances as compared to a Gauss-Newton (GN) approach while taking only a fraction of its computational time. It also estimates the covariance of the estimation errors which is necessary for our inlier test (see section 4.2). Recovering those covariances from the solution of a GN is computationally very expensive.

- *performing loop voting in order to infer the set of inliers.* In [23], the maximum loop length is limited to 6 in order not to take into account the uncertainty induced by the length of a loop. On the contrary, we explicitly model this uncertainty with the LG-IEKF and derive a statistical inlier test on matrix Lie groups. Thus we are able to close long loops. Consequently, even in the case of statistically independent outliers assumed in [23], our approach outperforms the algorithm proposed in [23].

Our approach works as follows: first of all, we sample an MST from the WAM $\mathcal{G}$. Then an LG-IEKF is applied on this MST with inlier test (see section 4.2) and loop voting (see section 4.2) at each loop closure. If every relative motion

**Algorithm 1** Robust approach

Inputs: $\mathcal{G}$ (weighted adjacency matrix), $\{Z_{ij}\}_{1\leq i<j\leq N}$ (relative motions), $\{\Sigma_{ii}\}_{1\leq i<j\leq N}$ (covariance matrices), $thresh$ (p-value of $\chi^2(p)$)

Outputs: $\mu$ (global motions), $P$ (full covariance matrix of global motions)

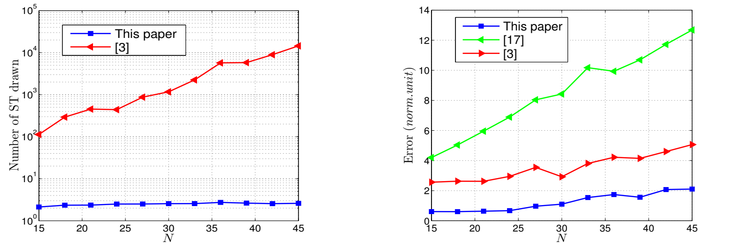
1. Draw an MST in $\mathcal{G}$ to get $\mathcal{E}_\mathcal{T} = (C^m)_{m=0:N-2}$, $\mathcal{L} = (M^m)_{m=1:N-2}$ and the tables of indices $i_\mathcal{T}$, $j_\mathcal{T}$, $i_\mathcal{L}$ and $j_\mathcal{L}$
2. Initialize[1] $\mu$ and $P$
   with inputs $C^0$, $Q_0 = \Sigma_{i_\mathcal{T}(0)i_\mathcal{T}(0)}$, $i_\mathcal{T}$, $j_\mathcal{T}$
   to get $\mu^{0|0}$ and $P^{0|0}$
3. Propagate[1] $\mu^{k-1|k-1}$ and $P^{k-1|k-1}$
   with inputs $\mu^{k-1|k-1}$, $P^{k-1|k-1}$, $C^k$, $Q_k = \Sigma_{i_\mathcal{T}(k)i_\mathcal{T}(k)}$, $j_\mathcal{T}$
   to get $\mu^{k|k-1}$ and $P^{k|k-1}$
4. Verify Loop Closures $M^k_\mathfrak{z}$ as explained in section 4.2
   with inputs $\mu^{k-1|k-1}_{i_\mathcal{L}(k,\mathfrak{z})R}$, $P^{k-1|k-1}_{i_\mathcal{L}(k,\mathfrak{z})i_\mathcal{L}(k,\mathfrak{z})}$, $\mu^{k-1|k-1}_{j_\mathcal{L}(k,\mathfrak{z})R}$, $P^{k-1|k-1}_{j_\mathcal{L}(k,\mathfrak{z})j_\mathcal{L}(k,\mathfrak{z})}$, $P^{k-1|k-1}_{i_\mathcal{L}(k,\mathfrak{z})j_\mathcal{L}(k,\mathfrak{z})}$, $M^k_\mathfrak{z}$, $\Sigma_{i_\mathcal{L}(k,\mathfrak{z})i_\mathcal{L}(k,\mathfrak{z})}$, $thresh$
   (a) if $M^k_\mathfrak{z}$ is not validated, remove it from $M^k$, $i_\mathcal{L}$ and $j_\mathcal{L}$
   (b) else mark as "checked" the path in $\mathcal{E}_\mathcal{T}$ that led to this loop closure as explained in section 4.2
5. Update[1] $\mu^{k|k-1}$ and $P^{k|k-1}$
   with inputs $\mu^{k|k-1}$, $P^{k|k-1}$, $M^k$, $R_k = \text{blkdiag}\left(\left\{\Sigma_{i_\mathcal{L}(k,\mathfrak{z})i_\mathcal{L}(k,\mathfrak{z})}\right\}_\mathfrak{z}\right)$, $i_\mathcal{L}$, $j_\mathcal{L}$
   to get $\mu^{k|k}$ and $P^{k|k}$
6. Iterate Step 3 to Step 6 until $k = N-2$
7. If every relative motion in $\mathcal{E}_\mathcal{T}$ has been "checked" at least once, return $\mu$ and $P$, otherwise up-weight (or remove) from $\mathcal{G}$ the relative motions involved in $\mathcal{E}_\mathcal{T}$ that have not been "checked" and go to 1.

[1] Further details on steps 2, 3 and 5 are provided as supplementary material.

that is part of the MST have been "checked" at least once, i.e each relative motion in the MST is involved in at least one validated loop closure, then the global motion estimates correspond to the output of the LG-IEKF. Otherwise, the relative motions involved in the MST that have not been "checked" are up-weighted (or deleted) from $\mathcal{G}$ and a new MST is drawn. The algorithm iterates until convergence of $\mathcal{G}$, i.e until an MST is completely checked. The complete algorithm is summarized in Algorithm 1.

**Limitations** The proposed approach may fail in several cases.

First of all, we assume that an MST of the WAM does not contain dependent outliers. If it does, depending on the graph structure, it may be possible to close loops that involve these dependent outliers. In this case, those outlier measurements are not rejected and the global motions are not correctly recovered. In practice, this limitation is not as strong as it appears. For instance, this was satisfied for all experiments shown in [3] (see the ground truth matrix of




(a) Number of ST drawn to find an ST without outliers.

(b) Estimation error of the global motions (please see text for details)

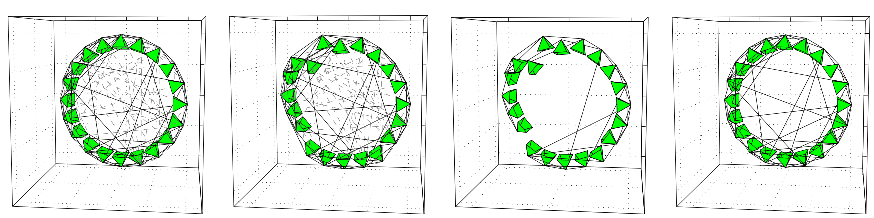
Fig. 2: Comparison of our approach to [3] and [18] on a camera pose estimation problem ($\lambda = \frac{1}{10}$ and number of relative motions fixed to $5N + n$ with $n = 60$).

each dataset in [3]) as well those shown in this paper (see section 5). Note that if there is only statistically independent outliers, our approach does not need a WAM to recover the global motions.

Secondly, if a relative motion measurement is deleted after an iteration of our algorithm, the graph might become disconnected. If it happens, it means that there is not enough redundancy in the graph structure to correctly recover the global motions. In this case, our approach can be applied to each connected component separately.

## 5 Applications and Results

In this section, the proposed framework is experimentally validated both on simulated and real data.



(a) True trajectory (b) Result of [18] (c) Result of [3] (d) This paper

Fig. 3: Camera Pose registration problem results, a cone represents a camera pose, a black line is an inlier measurement and a gray dashed line is an outlier

*Simulated data with independent outliers: Camera pose registration problem* In this section, we compare the performance of the proposed approach to two state of the art algorithms [3] and [18] on a camera pose registration problem (Lie group $SE(3)$). [18] was developed to deal with $SO(3)$ but its extension to $SE(3)$ is straightforward. We simulate circular camera trajectories (see Fig.3) with $N$ cameras where each camera $X_{iR_{True}}$ has a timestamp $t_i$ and we generate noisy relative motions as follows: first of all, a measurement can be either an inlier or an outlier. We model the probability of a measurement as $P(Z_{ij} \text{ is inlier}) = exp(-\lambda |t_i - t_j|)$ where $\lambda$ is a user-chosen parameter, i.e a larger time difference increases the chance to produce an outlier. After having drawn the label of a measurement (inlier or outlier), we sample the measurement. The distribution of the independent outliers is modeled as a centered Gaussian distribution on Lie groups with a large covariance matrix (the large covariance is not a problem in this case since $log_{SE(3)}^{\vee}$ is defined on the whole group) while an inlier can be sampled using (3). In our implementation, we use the Lie algebra basis of $\mathfrak{se}(3)$ given in [8] and the WAM $\mathcal{G}$ is built from the absolute time differences of the camera timestamps, i.e a low weight corresponds to a confident measurement and an infinite weight is given when a measurement is missing.

In Fig.2a, we show that the method proposed in [3] can be applied only on very small problems. Indeed, one can see that when the number of cameras increases ($N$ grows), it becomes quickly very difficult to draw an ST without outliers even with the help of a weighted adjacency matrix. In comparison, our approach, that iteratively updates the weighted adjacency matrix, always finds an ST without outliers in a few samplings.

In Fig.2b, we compare our optimization method (LG-IEKF with inlier test) against the robust approach proposed in [18] and the Expectation Maximization algorithm (EM) of [3]. The three approaches are provided the same outlier free spanning tree. [18] and [3] are initialized by composing the relative measurements of the ST as it is proposed by the authors of those papers. In order to compare the results of each approach to the true global motions, we need to add a step to align the estimated global motions with the true global motions. For that purpose, we apply a Gauss-Newton algorithm to minimize the sum of the following error: $\left\| log_{SE(3)}\left(\mu_{iR} X_{RR_{True}} X_{iR_{True}}^{-1}\right) \right\|^2$. The error obtained, for each approach, at convergence of the Gauss-Newton is presented in Fig.2b. We show that our method outperforms both [18] and [3]. Indeed, [18] is based on a robust convex L2-L1 norm to mitigate the influence of the outliers. However, because of the Lie group curvature, the complete functional is not convex. Therefore, the algorithm usually gets stuck in a poor local minimum. [3] introduces latent variables to classify the relative motions as inliers or outliers. However, the labels obtained at the initialization of the global motions are very difficult to modify. Indeed, the E-step does not take into account the estimation errors of the current global motion estimates which is negligible only when $N$ is small. Therefore, a lot of inlier relative motion remain classified as outliers. In comparison, our approach incrementally rejects outliers, taking into account the current uncertainty of the global motions, and refines its estimates with the inliers. Consequently, the global motions are correctly recovered. An example of recovered global motions with the three different approaches is presented Fig.3.

*Real data with statistically dependent outliers: Partial 3D reconstruction merging problem* The algorithm presented in this paper is applied to a partial 3D reconstruction merging problem (Lie group $Sim\,(3)$). Due to the lack of space, the details and results of this experiment are provided as supplementary material.

*Real data with statistically dependent outliers: Automatic planar image mosaicking problem* The algorithm presented in this paper is applied to an automatic planar image mosaicking problem. We took 53 photos of a planar scene (see Fig.4a) with a smartphone, detected points of interest and estimated the homographies (lie group $SL\,(3)$) between every pair of images using a RANSAC algorithm followed by a Gauss-Newton algorithm on $SL\,(3)$. The covariance matrix of each relative motion is obtained by inverting the approximated Hessian matrix once the Gauss-Newton has converged. In our implementation, we use the Lie algebra basis of $\mathfrak{sl}\,(3)$ given in [36] and the weighted adjacency matrix is obtained from the absolute time differences from the images timestamps. In this dataset, there are 65% of statistically dependent outliers (see Fig.4b) due to the ambiguity of the scene (some paper sheets are almost identical). In Fig.4, we compare the results of our approach against the EM algorithm of [3] which is initialized by composing the relative homographies of the MST obtained with our algorithm. On the one hand, once again, the proposed EM of [3] classifies a lot of inliers as outliers since the estimation errors of the global motions estimates is not taken into account during the E-step. Consequently, [3] is not able to correctly recover the global motions (see Fig.4c). On the other hand, our approach perfectly infers the set of inliers and produces a very precise mosaic (see Fig.4d). We could not apply [18] because $log_{SL(3)}$ is not defined on the whole group. We also tried to compare our formalism to the openCV implementation of [5], however, due to the high ambiguity of the scenes, it was not able to produce any result.

## 6 Discussion & Conclusion

First of all, we would like to stress 3 aspects concerning the paper:
- This work addresses the fundamental problem of robust motion averaging for any Lie group especially "mixed" groups such as $SE\,(3)$, $SL\,(3)$ or $Sim\,(3)$;
- It deals with cases where there are more correlated outliers than inliers. For instance, our image mosaicking dataset has 65% of dependent outliers. Therefore, we are beyond the stage at which the independent outlier assumption starts to degrade gracefully (see comparison with [23] in [3]);
- The proposed approach significantly outperformed the two state of the art algorithms [3] and [18], on $SE\,(3)$, $SL\,(3)$ and $Sim\,(3)$.

The contributions[2] of the paper are:
- Proper handling of non-isotropic covariances on Lie groups coupled with an

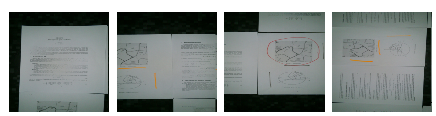
[2] The supplementary material and the Matlab code are available at https://sites.google.com/site/guillaumebourmaud/

efficient incremental approach to avoid local minima;
- Definition of a new $\chi^2$ inlier test that deals with "mixed" groups;
- A new tree sampling scheme aimed at significantly reducing the computational cost of the sampling scheme of [3]. This new sampling scheme allows to handle a much larger number of motions $N$ (typically $N = 1000$). $N$ is no longer limited by the sampling scheme but only by the memory size of the estimated covariance matrix ($6N \times 6N$ for $SE(3)$);

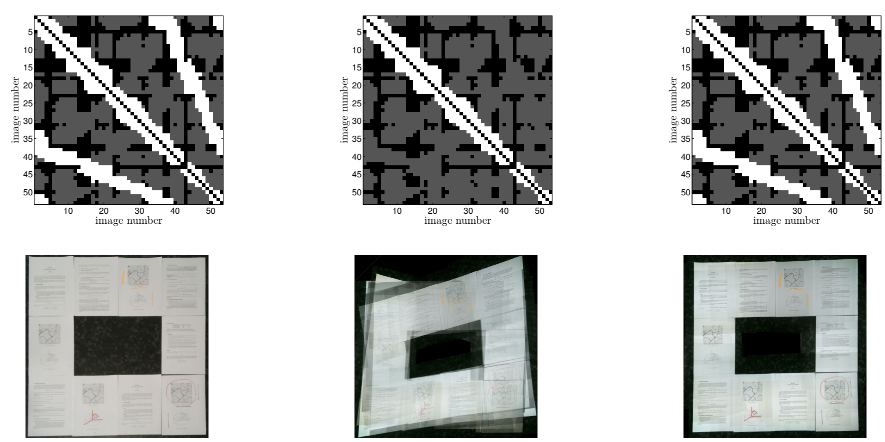
The first 3 contributions of this paper could be applied inside a Structure from Motion pipeline such as [3] or [21] but this is out of the scope of the current paper and left as future work. Thus, when comparing to [3], we consider only their robust motion averaging approach and did not use their Structure from Motion datasets. Consequently, we compared neither to [21] nor to [24] (that do not solve any motion averaging problem).

In fact, our contributions concern mainly "mixed" groups which do not have robust enough solutions yet. The proposed approach is based on a generic matrix Lie group formulation, which should be usable on a wide variety of applications.

**Acknowledgment** The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement 288199 - Dem@Care. The authors would like to thank the reviewers, Moncef Hidane and Cornelia Vacar for their valuable help.



(a) Examples of input images




(b) Ground Truth (c) [3] (d) This paper

Fig. 4: Image mosaicking: in the labeling matrices a white pixel is an inlier, a black pixel corresponds to an unavailable measurement, a gray pixel corresponds to an outlier. Observe that our labeling inlier/outlier is perfect (see 4d).

## References


1. Agrawal, M.: A Lie algebraic approach for consistent pose registration for general Euclidean motion. In: IROS. (2006) 1891–1897
2. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM **24** (1981) 381–395
3. Roberts, R., Sinha, S.N., Szeliski, R., Steedly, D.: Structure from motion for scenes with large duplicate structures. In: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, IEEE (2011) 3137–3144
4. Hartley, R., Trumpf, J., Dai, Y., Li, H.: Rotation averaging. International Journal of Computer Vision (2013) 1–39
5. Brown, M., Lowe, D.G.: Automatic panoramic image stitching using invariant features. International Journal of Computer Vision **74** (2007) 59–73
6. Wachinger, C., Wein, W., Navab, N.: Registration strategies and similarity measures for three-dimensional ultrasound mosaicing. Academic radiology **15** (2008) 1404–1415
7. Chirikjian, G.S.: Stochastic Models, Information Theory, and Lie Groups, Volume 2. Springer-Verlag (2012)
8. Selig, J.M.: Lie groups and Lie algebras in robotics. NATO Science Series II: Mathematics, Physics and Chemistry (2005) 101–125
9. Malis, E., T.Hamel, Mahony, R., Morin, P.: Dynamic estimation of homography transformations on the special linear group of visual servo control. In: IEEE Conference on Robotics and Automation. (2009)
10. Strasdat, H., Davison, A., Montiel, J., Konolige, K.: Double window optimisation for constant time visual SLAM. In: IEEE International Conference on Computer Vision (ICCV). (2011)
11. Huber, P.J., et al.: Robust estimation of a location parameter. The Annals of Mathematical Statistics **35** (1964) 73–101
12. Bourmaud, G., Mégret, R., Giremus, A., Berthoumieu, Y.: Global motion estimation from relative measurements using iterated extended Kalman filter on matrix Lie groups. (Accepted to ICIP 2014)
13. Bandeira, A.S., Singer, A., Spielman, D.A.: A Cheeger inequality for the graph connection laplacian. arXiv preprint arXiv:1204.3873 (2012)
14. Singer, A., Shkolnisky, Y.: Three-dimensional structure determination from common lines in cryo-em by eigenvectors and semidefinite programming. SIAM journal on imaging sciences **4** (2011) 543–572
15. Boumal, N., Singer, A., Absil, P.A.: Robust estimation of rotations from relative measurements by maximum likelihood. In: Proceedings of the 52nd Conference on Decision and Control, CDC. (2013)
16. Wang, L., Singer, A.: Exact and stable recovery of rotations for robust synchronization. Information and Inference **2** (2013) 145–193
17. Hartley, R., Aftab, K., Trumpf, J. In: L1 rotation averaging using the Weiszfeld algorithm. Institute of Electrical and Electronics Engineers (2011) 3041–3048
18. Chatterjee, A., Govindu, V.M.: Efficient and robust large-scale rotation averaging. In: ICCV. (2013)
19. Crandall, D.J., Owens, A., Snavely, N., Huttenlocher, D.: Discrete-continuous optimization for large-scale structure from motion. In: CVPR. (2011) 3001–3008
20. Enqvist, O., Kahl, F., Olsson, C.: Non-sequential structure from motion. In: ICCV Workshops. (2011) 264–271

21. Jiang, N., Cui, Z., Tan, P.: A global linear method for camera pose registration. (2013 IEEE International Conference on Computer Vision)
22. Moulon, P., Monasse, P., Marlet, R.: Global fusion of relative motions for robust, accurate and scalable structure from motion. (2013 IEEE International Conference on Computer Vision)
23. Zach, C., Klopschitz, M., Pollefeys, M.: Disambiguating visual relations using loop constraints. In: Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, IEEE (2010) 1426–1433
24. Jiang, N., Tan, P., Cheong, L.F.: Seeing double without confusion: Structure-from-motion in highly ambiguous scenes. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE (2012) 1458–1465
25. Govindu, V.M.: Robustness in motion averaging. In: Computer Vision–ACCV 2006. Springer (2006) 457–466
26. Wilson, K., Snavely, N.: Network principles for sfm: Disambiguating repeated structures with local context. (2013 IEEE International Conference on Computer Vision)
27. Sunderhauf, N., P.Protzel: Switchable constraints for robust pose graph slam. In: IROS. (2012)
28. Latif, Y., Cadena, C., Neira, J.: Robust loop closing over time for pose graph SLAM. The International Journal of Robotics Research (2013)
29. Olson, E., Agarwal, P.: Inference on networks of mixtures for robust robot mapping. In: RSS. (2012)
30. Bourmaud, G., Mégret, R., Giremus, A., Berthoumieu, Y.: Discrete extended Kalman filter on Lie groups. In: Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European. (2013)
31. Wang, Y., Chirikjian, G.: Error propagation on the Euclidean group with applications to manipulators kinematics. IEEE Transactions on Robotics **22** (2006)
32. Wolfe, K., Mashner, M., Chirikjian, G.: Bayesian fusion on Lie groups. Journal of Algebraic Statistics **2** (2011) 75–97
33. Barfoot, T.D., Furgale, P.T.: Associating uncertainty with three-dimensional poses for use in estimation problems. IEEE Trans. Robot. (**30**) 679–693
34. Fisher, R.A., Yates, F., et al.: Statistical tables for biological, agricultural and medical research. Statistical tables for biological, agricultural and medical research. (1949)
35. Scheinerman, E.R.: (Matgraph: a matlab toolbox for graph theory)
36. Benhimane, S., Malis, E.: Homography-based 2d visual tracking and servoing. The International Journal of Robotics Research **26** (2007) 661–676