

CONVEX IMAGE ORIENTATION FROM RELATIVE ORIENTATIONS

M. Reich*, C. Heipke

Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, Nienburger Str. 1, D-30167 Hannover, Germany
reich@ipi.uni-hannover.de, heipke@ipi.uni-hannover.de

Commission III, WG III/1

KEY WORDS: image orientation, pose estimation, rotation averaging, Lie algebra, structure-from-motion, spatial intersection

ABSTRACT:

In this paper we propose a novel workflow for the estimation of global image orientations given relative orientations between pairs of overlapping images. Our approach is convex and independent on initial values. First, global rotations are estimated in a relaxed semidefinite program (SDP) and refined in an iterative least squares adjustment in the tangent space of $SO(3)$. A critical aspect is the handling of outliers in the relative orientations. We present a novel heuristic graph based approach for filtering the relative rotations that outperforms state-of-the-art robust rotation averaging algorithms. In a second part we make use of point-observations, tracked over a set of overlapping images and formulate a linear homogeneous system of equations to transfer the scale information between triplets of images, using estimated global rotations and relative translation directions. The final step consists of refining the orientation parameters in a robust bundle adjustment. The proposed approach handles outliers in the homologous points and relative orientations in every step of the processing chain. We demonstrate the robustness of the procedure on synthetic data. Moreover, the performance of our approach is illustrated on real world benchmark data.

1 INTRODUCTION

Image orientation (or structure-from-motion, pose estimation) comprises the computation of the global translation and rotation in a single coordinate system for every element of a set of overlapping images. Although this is a well understood problem in photogrammetry, it gained recent interest with the increasing amount of images available, e.g. from photo-sharing websites (Pollefeys et al., 2004; Agarwal et al., 2009; Wu et al., 2011). Moreover, novel global models evolved recently thanks to increasing computational power and recent research in optimization techniques (Martinec and Pajdla, 2007; Arie-Nachimson et al., 2012; Sinha et al., 2012; Jiang et al., 2013; Moulon et al., 2013; Özyesil et al., 2015). These models use all relative orientations between pairs or triplets of images at once (which is the reason for the term *global* in this context) to often subsequently estimate global rotations and translations. In contrast to sequential models, which start with a small set of images, append further images and refine the solution with intermediate bundle adjustment to reduce drift, global models often do not suffer from the arbitrary selection of the initial set and the order in which images are appended. However, while achieving impressive results, many global approaches tend to over-relax the problem of estimating rotations in terms of omitting the orthonormality constraint (Martinec and Pajdla, 2007; Sinha et al., 2012; Jiang et al., 2013; Moulon et al., 2013) or suffer from degeneracy for collinear camera composition (Govindu, 2001; Arie-Nachimson et al., 2012).

In this paper we present a novel convex global image orientation approach that primarily focuses on the accuracy of the parameters to be estimated (Fig. 1). The main contribution of this paper is *twofold*: *First*, we propose a novel combination of a robust, convex and maximum likelihood estimation of global rotations considering covariance information propagated from homologous points. *Second*, we extend the linear translation

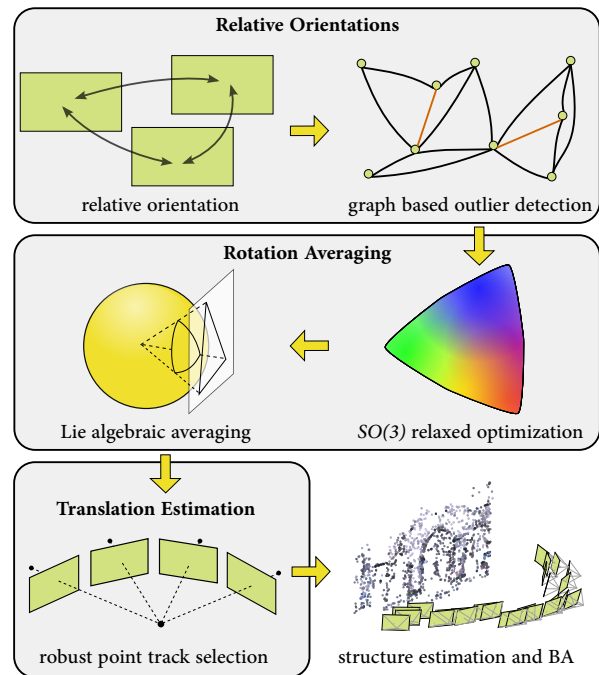


Figure 1. Workflow of our image orientation model.

estimation algorithm of Cui et al. (2015) and add a simple outlier detection strategy using local spatial intersection.

We evaluate the performance of the proposed approach with respect to its robustness using synthetic data. Further experiments on benchmark data demonstrate the value of our work regarding its accuracy for the derived orientations compared to existing approaches.

*Corresponding author

2 RELATED WORK

Research in the field of global image orientation is currently very active. Often, the problem is visualized by a graph in which the vertices represent the unknown global orientations that are connected by an edge if a relative orientation has been estimated. In recent years many works were published that, while pursuing different algorithms, address the problem by estimating rotations and translations separately (Martinec and Pajdla, 2007; Arie-Nachimson et al., 2012; Sinha et al., 2012; Jiang et al., 2013; Moulon et al., 2013; Arrigoni et al., 2015). In the following we first review related work on estimating rotations followed by an overview of existing work towards translation estimation.

Rotation estimation: One of the first publications comprising the estimation of global rotations from relative estimates is the fundamental work of Govindu (2001) in which rotations represented as quaternions are estimated in a least squares sense without applying a unit or sign constraint. A few years later, inspired by developments in group theory (e.g. (Kanatani, 1990)), Govindu (2004) proposed a novel model that iteratively and jointly estimates rotations and translations in their respective Lie algebra. Martinec and Pajdla (2007) propose a least squares estimation of rotations in which the orthonormality is enforced by a subsequent mapping to the nearest member of $SO(3)$ using the Frobenius norm. It is also noted that an unconstrained optimization using quaternions as in Govindu (2001) should not be used in practice because the optimal solution can be far from a valid quaternion regarding the unit-norm constraint. In Fredriksson and Olsson (2012), who propose a SDP (i.e. the dual problem) to provide a lower bound for the optimal rotations, it is shown that the local linear solution from Govindu (2001) including a single norm constraint for all quaternions in most cases leads to the global optimal solution. The work of Arie-Nachimson et al. (2012) extends the linear model of Martinec and Pajdla (2007) and proposes a spectral method that is more robust and can be cast into a SDP that comprises an orthogonality constraint.

In our model we go one step further and extend the SDP by a convex constraint that requires the resulting rotation matrices to be in the convex hull of $SO(3)$, yielding a compatible linear matrix inequality constraint. Subsequently, we refine the rotations in an iterative least squares estimation similar to Govindu (2004), but including covariance information for the relative rotations as prior. Although rotations are estimated iteratively, the algorithm is efficient and has the advantage that it is in $SO(3)$, respectively $so(3)$, hence no subsequent adaptation of rotations is needed.

For enhancing robustness Hartley et al. (2011, 2013) proposed an iterative implementation of the Weiszfeld algorithm (Weiszfeld and Plastria, 2009) that minimizes the L_1 norm in sequential graph optimization. Chatterjee and Govindu (2013) present a two-step procedure, first comprising a L_1 -norm averaging and then a M-estimation. Similar to the model presented in this work Enqvist et al. (2011) try to find outliers using cycle errors in the view graph starting from a most reliant minimum spanning tree (MST) and then adding further edges step by step. While the approach is not limited by a cycle length as in Zach et al. (2010) several search heuristics have to be applied to cover the case the initial MST is affected by outliers.

Our approach is similar to Zach et al. (2010); Enqvist et al. (2011) but is based on a breadth-search propagation through the view graph.

Translation estimation: Recent related work in the field of translation estimation with known rotations can be roughly

divided into two categories: First, there are quasiconvex L_∞ -norm optimizations solving a second-order-cone program (Hartley and Schaffalitzky, 2004; Ke and Kanade, 2007; Martinec and Pajdla, 2007; Kahl and Hartley, 2008; Olsson and Kahl, 2010) in which translations and object points are estimated simultaneously. These approaches allow a globally optimal solution at the price of high outlier sensitivity and computational inefficiency for large problems. Second, models were suggested comprising the relative translation direction that treat orientations and structure separately (Govindu, 2001, 2004; Jiang et al., 2013; Moulon et al., 2013; Özyesil and Singer, 2015; Cui et al., 2015). Our model belongs to the second category.

Linear models founded on pairwise relative translation directions (Govindu, 2001; Arie-Nachimson et al., 2012) can be solved efficiently but easily fall victim to degeneracy for collinear camera composition. Triplet constraint approaches (Jiang et al., 2013; Moulon et al., 2013) exhibit stronger resistance to this degeneracy but depend on a highly connected triplet graph. In Jiang et al. (2013) a geometric error is minimized based on a linear formulation of a “triangulation” of a third camera location from two relative directions. Moulon et al. (2013) minimize the L_∞ -norm of relative Euclidean distances in a linear program. The work of Özyesil and Singer (2015) minimizes the same objective function applying a more robust L_1 -norm formulated in an iterative least squares algorithm. To be robust against outliers in homologous points they propose a pairwise direction estimation using averaged rotations instead of relative estimates. In Özyesil et al. (2015) global translations are recovered from relative directions using semidefinite programming enabling constraints that avoid a clustering of the solution.

We adapt the model of Cui et al. (2015) using tracked point observations to transfer the scale between images. The strength of this model lies in the ability of processing images with weak overlap. We extend their model by a simple outlier detection via local spatial intersection for points that survived the robust estimation of relative orientations.

The remainder of this paper is structured as follows: In Section 3 we describe the estimation of relative orientations. A focus lies on a formal definition of our novel graph based outlier detection algorithm. Section 4 is dedicated to rotation averaging. First, our convex SDP formulation is examined followed by an insight into our iterative refinement model. The estimation of global translations is developed in Section 5. Our model is evaluated in section 6 and finally conclusions are given in Section 7.

3 ESTIMATION OF RELATIVE ORIENTATIONS

Our model is founded on a set of homologous points for a set of calibrated images. We assume this set to include outliers which is a logical assumption considering practical applications. Hence, besides estimating precise pairwise relative orientation parameters, the focus lies on first providing an outlier-free set of homologous points and subsequently on detecting inconsistencies in the estimated relative orientations. As an initial step we deploy a robust estimation of the essential matrix in a RANSAC estimation (Fischler and Bolles, 1981). While many outliers are rejected during this procedure, some may accidentally fulfill the epipolar constraint and remain in the set of homologous points. One idea is to use the more robust trifocal tensor for relative orientations (Jiang et al., 2013; Moulon et al., 2013). However, we stick to image pairs and apply a subsequent constrained iterative least-squares adjustment of the relative orientation parameters to detect remaining outliers in the correspondences. We see two benefits of this approach: First, the orientations

are refined leading to a superior overall accuracy and second, full covariance information for every relative orientation can be extracted which is useful for the following computations. Of course, our approach is also applicable to pairwise relative orientations from trifocal tensors.

3.1 Constrained Iterative Optimization

Let $\mathbf{R}_{ij}^0 \in SO(3)$ and $\mathbf{t}_{ij}^0 \in \mathbb{R}^3$ be the relative rotation and translation for the image pair $\{i, j\}$ with $(i, j) \in \{1, \dots, n\}, i \neq j$ and a set of n images. We want to refine these values for every image pair $\{i, j\}$ using the nonlinear residual function

$$\mathbf{v}_{ij} = \mathbf{m}_j^T \mathbf{R}_{ij}^T [\mathbf{t}_{ij}]_{\times} \mathbf{m}_i, \quad (1)$$

with $\mathbf{m}_{i,j} = \mathbf{K}_{i,j}^{-1} \mathbf{x}_{i,j}$ being the image coordinates of the homologous points normalized with the inverse calibration matrix $\mathbf{K}_{i,j}^{-1}$ and $[\cdot]_{\times}$ mapping a vector to a skew-symmetric matrix. While this comprises the estimation of six parameters, only five of them are linearly independent. We propose the following optimization model:

$$\begin{aligned} & \underset{\mathbf{R}_{ij}, \mathbf{t}_{ij}}{\text{minimize}} && \mathbf{v}_{ij}^T \mathbf{v}_{ij} \\ & \text{subject to} && \mathbf{R}_{ij} \in SO(3), \quad |\mathbf{t}_{ij}| = 1. \end{aligned} \quad (2)$$

Problem (2) is solved in a Gauss-Helmert model by linearizing Eq. (1) using \mathbf{R}_{ij}^0 and \mathbf{t}_{ij}^0 as initial values. The parameters are improved iteratively until convergence. Observations are weighted based on their variances (see Förstner et al. (2004), p. 813) which allows for a final elimination of remaining outliers. Lastly, we derive the covariance matrices for each relative orientation $\Sigma_{\mathbf{R}_{ij}}$ and $\Sigma_{\mathbf{t}_{ij}}$.

3.2 Graph Optimization

The goal of graph optimization is the detection of defective relative orientations. Let \mathcal{E} denote the set of all relative orientations, estimated as described in the previous section. Then, we can formulate the undirected graph $G = (\mathcal{V}, \mathcal{E})$ in which \mathcal{V} corresponds to the set of global orientations $\{\mathbf{R}_i, \mathbf{t}_i\}$. Every pair of vertices in \mathcal{V} is connected by an edge if a relative orientation between the corresponding images has been estimated. In general, the number of edges m in G is higher than the number of vertices n which induces redundancy, a prerequisite for outlier detection. We propose a sequential graph propagation model that uses the concatenation constraint of rotation matrices $\mathbf{R}_j = \mathbf{R}_i \mathbf{R}_{ij}$ to find defective relative orientations. In the following we concentrate on rotations and use \mathcal{E} and \mathcal{V} representing a set of relative and global rotations, respectively, with a graph structure equivalent to G .

The first starting vertex $\mathcal{V}_{i,0}$, without loss of generality, is the vertex with maximum degree. Successively, global rotations are propagated to adjacent vertices adopting a different vertex as starting point in every sequence. In the following, the starting point for the iteration $t \in \{1, \dots, n-1\}$ is the vertex whose rotation was propagated by the maximum number of adjacent vertices: $\mathcal{V}_{i,t} = \mathcal{V}_{\max|\mathcal{E}_{i,t-1}|}$, with $\mathcal{E}_{i,t-1} \subseteq \mathcal{E}_i \subseteq \mathcal{E}$ being the set of edges used for the propagation of \mathcal{V}_i at sequence $t-1$.

If, during propagation from i to j , a vertex \mathcal{V}_j already exhibits an estimated rotation we calculate the angular similarity (Hartley et al., 2013) of both proposals $\{\mathcal{V}_{j,t}, \mathcal{V}_{i,t} \mathcal{E}_{ij}\}$ and decide, respecting a predefined threshold τ_s , whether both rotations coincide or one estimation is affected by an outlier. In the former case both estimations are averaged using single rotation averaging (Hartley et al., 2013) weighted with the covariance information.

In the latter case we consult the whole set of edges \mathcal{E}_j connected with the respective vertex \mathcal{V}_j and consider all already estimated vertices to make an estimation for \mathcal{V}_j . A new estimation is averaged from the largest subset of compliant rotation matrices $\mathcal{E}_j^+ \subseteq \mathcal{E}_j$ (matrices, which lead to rotations whose differences are smaller than τ_s). With differing relative rotations $\mathcal{E}_j^- = \mathcal{E}_j \setminus \mathcal{E}_j^+$, all members of \mathcal{E}_j^- are classified as outliers if $|\mathcal{E}_j^+|/|\mathcal{E}_j^-| > \tau_c$, with τ_c being another threshold. Finally, after n sequences, all relative rotations are checked again if they support the estimation which leads to a final classification into \mathcal{E}^+ and \mathcal{E}^- . A pseudo code for the graph optimization is provided in Algorithm 1.

Our model is based on the assumption that there exists a redundant number of relative rotations used to estimate global rotations. Moreover, we argue that outliers are uniformly distributed, i.e. it is unlikely that there will be a consistent solution using only outliers for the estimation of a rotation. Given these constraints, our algorithm will converge (i.e. it behaves like a breadth-search) and we end up with a set $\mathcal{E}^- \subseteq \mathcal{E}$ defining deficient relative rotations (outliers). In the following only relative rotations in \mathcal{E}^+ are consulted.

Algorithm 1 Graph Optimization

```

1: initialization:  $\mathcal{V}_{i,0}, \mathcal{E}^- = \emptyset$ 
2: propagation from  $\mathcal{V}_{i,0}$ 
3: for every  $\mathcal{V}_{i,t}, t \in \{1, \dots, n-1\}$  do
4:   for every  $\mathcal{E}_{ij} \in \mathcal{E}_i$  do
5:     propagation from  $\mathcal{V}_{i,t}$ 
6:     if cycle closed then
7:       if  $\text{sim}(\mathcal{V}_{j,t}, \mathcal{V}_{i,t} \mathcal{E}_{ij}) \leq \tau_s$  then
8:          $\mathcal{V}_{j,t} \leftarrow \text{mean}(\mathcal{V}_{j,t}, \mathcal{V}_{i,t} \mathcal{E}_{ij})$ 
9:       else
10:        find  $\mathcal{E}_j^+ \subseteq \mathcal{E}_j$ 
11:         $\mathcal{V}_{j,t} \leftarrow \text{mean}(\mathcal{V}_{k,t} \mathcal{E}_{kj}), \forall \mathcal{E}_{kj} \in \mathcal{E}_j^+$ 
12:        if  $|\mathcal{E}_j^+|/|\mathcal{E}_j^-| > \tau_c$  then
13:           $\mathcal{E}^- \leftarrow \{\mathcal{E}^-, \mathcal{E}_j^-\}$ 
14:         $\mathcal{V}_{i,t+1} \leftarrow \mathcal{V}_{\max|\mathcal{E}_{i,t-1}|}$ 
15: for every  $\{i, j\} \in \{1, \dots, n\}$  do
16:   check  $\text{sim}(\mathcal{E}_{ij}, \mathcal{V}_i^T \mathcal{V}_j)$  and update  $\{\mathcal{E}^+, \mathcal{E}^-\}$ 
17:  $\mathcal{E}^+ \leftarrow \mathcal{E} \setminus \mathcal{E}^-$ 

```

4 ROTATION AVERAGING

In the previous section we demonstrated the derivation of a consistent set of precise relative rotations \mathcal{E}^+ including full covariance information for each element in \mathcal{E}^+ . Our goal during rotation averaging is first to find an optimal set \mathcal{V}^* corresponding to global rotations and subsequently refine it to derive a set \mathcal{V}^{*+} . First, we describe a convex relaxed SDP for the generation of initial rotations (Sec. 4.1), followed by a demonstration of our Lie algebraic least squares refinement referring to (Govindu, 2004) (Sec. 4.2).

4.1 $SO(3)$ Relaxed Optimization

Like most image orientation or structure-from-motion approaches, e.g. (Martinec and Pajdla, 2007; Arie-Nachimson et al., 2012; Arrigoni et al., 2015), we aim at minimizing the cost function arising from the concatenation constraint of rotations:

$$\begin{aligned} & \underset{\{\mathbf{R}_1, \dots, \mathbf{R}_n\}}{\text{minimize}} && \sum_{(i,j)=1}^n \|\mathbf{R}_i^T \mathbf{R}_j - \mathbf{R}_{ij}\|_F \\ & \text{subject to} && \mathbf{R}_1, \dots, \mathbf{R}_n \in SO(3), \end{aligned} \quad (3)$$

with $\|\cdot\|_F$ being the Frobenius norm. The $SO(3)$ -constraint in (3) includes the nonlinear constraint $\det(\mathbf{R}_1, \dots, \mathbf{R}_n) = 1$. We are guided by recent work on optimization on special groups, e.g. (Horowitz et al., 2014; Saunderson et al., 2014). These authors demonstrate that the convex hull of $SO(3)$ can be described as a linear matrix inequality, yielding a valid SDP-formulation: a matrix \mathbf{R} is inside the convex hull of $SO(3)$, i.e. $\mathbf{R} \in \text{conv}(SO(3))$, if and only if

$$\begin{bmatrix} 1 - R_{11} - R_{22} + R_{33} & R_{13} + R_{31} & R_{12} - R_{21} & R_{23} + R_{32} \\ R_{13} + R_{31} & 1 + R_{11} - R_{22} - R_{33} & R_{23} - R_{32} & R_{12} + R_{21} \\ R_{12} - R_{21} & R_{23} - R_{32} & R_{12} + R_{21} & R_{23} + R_{32} \\ R_{23} + R_{32} & R_{12} + R_{21} & R_{23} + R_{32} & R_{12} - R_{21} \end{bmatrix} \succeq 0, \quad (4)$$

with $R_{row,col}$ denoting elements in \mathbf{R} and $[\cdot] \succeq 0$ meaning positive semi definiteness of a matrix.

In order to solve optimization problem (3) we define matrix \mathbf{M} being composed of a comprehensive pairwise multiplication of all unknown rotations, hence if $\mathbf{D} = [\mathbf{R}_1, \dots, \mathbf{R}_n]$, $\{\mathbf{R}_1, \dots, \mathbf{R}_n\} \in SO(3)$ then $\mathbf{M} = \mathbf{D}^T \mathbf{D}$,

$$\mathbf{M} = \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{R}_1^T \mathbf{R}_2 & \dots & \mathbf{R}_1^T \mathbf{R}_n \\ \mathbf{R}_2^T \mathbf{R}_1 & \mathbf{I}_{3 \times 3} & \dots & \mathbf{R}_2^T \mathbf{R}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_n^T \mathbf{R}_1 & \mathbf{R}_n^T \mathbf{R}_2 & \dots & \mathbf{I}_{3 \times 3} \end{bmatrix}. \quad (5)$$

Thus, \mathbf{M} is a symmetric matrix composed of 3×3 submatrices $\mathbf{R}_i^T \mathbf{R}_j$. We refer to these submatrices using indices i and j as row- and column-index, respectively. Our objective function is derived as a function of those submatrices \mathbf{M}_{ij} of \mathbf{M} , for which a corresponding relative rotation \mathbf{R}_{ij} has been estimated. From Eq. (6) we can derive the linear cost function $f(\mathbf{M}_{ij}) = -\text{tr}(\mathbf{M}_{ij} \mathbf{R}_{ij}^T)$ with $j \geq i$, hence using the upper triangular part of \mathbf{M} (see Appendix Eq. (17) for a derivation of the cost function). We derive the following SDP:

$$\begin{aligned} & \underset{\{\mathbf{R}_1, \dots, \mathbf{R}_n\}}{\text{minimize}} && \sum_{(i,j,k,l)=1}^n f(\mathbf{M}_{kl}) \\ & \text{subject to} && \mathbf{M} \succeq 0, \quad \mathbf{M}_{kk} = \mathbf{I}_{3 \times 3} \\ & && \mathbf{M}_{kl} \in \text{conv}(SO(3)). \end{aligned} \quad (6)$$

In Saunderson et al. (2014) it is empirically demonstrated on synthetic examples that this formulation often leads to rotations being part of $SO(3)$ and achieves superior results to $O(3)$ -based relaxations as used e.g. in Arie-Nachimson et al. (2012). It is important to note that (6) is not robust against outliers; we argue that all defective rotations have already been eliminated during graph optimization. Using the formulation in (6) we elegantly achieve an optimal set of absolute rotations \mathcal{V}^* by a singular value decomposition (SVD) of the optimal matrix \mathbf{M}^* that does neither depend on the order during propagation nor the initial starting vertex.

4.2 Lie Algebraic Averaging

In order to further refine the estimated rotations and find a set of improved rotations $\mathcal{V}^{*+} = \{\mathbf{R}_1^{*+} \dots \mathbf{R}_n^{*+}\}$ we perform a Lie algebraic least squares optimization using the rotations in \mathcal{V}^* as initialization and the covariance information for the relative rotations as prior information for weighting. We follow

in spirit the algorithm of Govindu (2004) but estimate the rotations only. Note that in principle the formulation allows an iterative reweighting to find remaining outliers. We assume outlier elimination has been done exhaustively during graph optimization (Sec. 3.2) and merely augment the influence of accurate relative rotations against weaker ones, i.e. finding a weighted average solution. We linearize the functional model $\Delta \mathbf{R}_{ij} = \mathbf{R}_i \mathbf{R}_{ij} \mathbf{R}_j^T$ using the matrix logarithm, $\log(\cdot) : SO(3) \rightarrow \mathfrak{so}(3)$, the first order approximation of the *Baker-Campbell-Hausdorff*-formula (Gilmore, 1974): $\log(\mathbf{X}\mathbf{Y}) \approx \log(\mathbf{X}) + \log(\mathbf{Y})$, $\mathbf{X}, \mathbf{Y} \in SO(3)$ and initial values $\mathbf{R}_1^*, \dots, \mathbf{R}_n^* \in \mathcal{V}^*$:

$$\log(\Delta \mathbf{R}_{ij}) = \log(\mathbf{R}_i^{*+} \mathbf{R}_i^{*T} \mathbf{R}_i^* \mathbf{R}_{ij} \mathbf{R}_j^{*T} \mathbf{R}_j^* \mathbf{R}_j^{*+T}) \quad (7)$$

$$\log(\Delta \mathbf{R}_{ij}) \approx \log(\mathbf{R}_i^{*+} \mathbf{R}_i^{*T}) + \log(\mathbf{R}_i^* \mathbf{R}_{ij} \mathbf{R}_j^{*T}) \dots - \log(\mathbf{R}_j^{*+} \mathbf{R}_j^{*T}). \quad (8)$$

The Lie algebra $\mathfrak{so}(3)$ comprises all skew-symmetric matrices in $\mathbb{R}^{3 \times 3}$. Let $[\cdot]_{\times}^{-1} : \mathbb{R}^{3 \times 3} \rightarrow \mathbb{R}^3$, $\mathbf{v}_{\Delta \mathbf{R}}$ be a $[3m \times 1]$ vector, stacked with $[\log(\Delta \mathbf{R}_{ij})]_{\times}^{-1}$, for all $\mathbf{R}_{ij} \in \mathcal{E}^+$ and let $\mathbf{r}_i^{*+} = [\log(\mathbf{R}_i^{*+} \mathbf{R}_i^{*T})]_{\times}^{-1}$. We then establish the following optimization problem:

$$\underset{\{\mathbf{r}_1^{*+}, \dots, \mathbf{r}_n^{*+}\}}{\text{minimize}} \quad \mathbf{v}_{\Delta \mathbf{R}}^T \mathbf{P} \mathbf{v}_{\Delta \mathbf{R}}. \quad (9)$$

\mathbf{P} is a $[3m \times 3m]$ weight matrix, computed from every individual covariance matrix for each \mathbf{R}_{ij} in order to control their influence on the solution $\{\mathbf{R}_1^{*+}, \dots, \mathbf{R}_n^{*+}\}$. Since the objective is linearized at the initial values, the optimization is performed iteratively until convergence taking the improved rotations as initialization for the subsequent iteration: $\mathbf{R}_i^{*+} = \exp([\mathbf{r}_i^{*+}]_{\times}) \mathbf{R}_i^*$. Note that the functional model is not affected by the update (the Jacobian consists of positive and negative unit matrices), which makes this local optimization efficient.

5 TRANSLATION ESTIMATION

In this section we describe how global translation parameters are estimated for every image in \mathcal{V} using the normalized translation directions \mathbf{t}_{ij}^0 , the global rotations in \mathcal{V}^{*+} and the observations of points that are visible in at least three overlapping images. We follow the approach developed in Cui et al. (2015) and use a minimal configuration of three images that share the observation of a single point in object space for which a linear constraint is defined. Instead of performing a L_1 optimization, we combine a simple outlier elimination and a maximum likelihood estimation of the unknown global translations. In Sec. 5.1 we review the construction of linear constraints from tracked observations whereas Sec. 5.2 is dedicated to our extension, i.e. the deployment of appropriate conditions for the observations to be considered an inlier.

5.1 Linear Constraints from Point Tracks

In most occurrences during pose estimation there will be object points that are observed in three or more images. For the ease of understanding let there be one object point \mathbf{X} which is observed in a set of images \mathcal{I} . The position of this point in object space can be estimated for every possible pairwise combination of images in \mathcal{I} via spatial intersection, given that a relative translation direction between the images is available. Of course, every reconstruction is valid only in a local coordinate system spanned by each pair of images. Without loss of generality this local

coordinate system is located at one of the image projection center and the scale is set according to a distance of 1 between the two projection centers. Since global rotations for the images are known, all local reconstructions are rotated consistently. With the assumption that all local reconstructions must meet the same position in object space, which we implicitly manifest by assuming that the observations in image space represent the same point, we can establish a constraint concerning the scale of individual relative translation directions \mathbf{t}_{ij}^0 .

More formally, using image i and j the point \mathbf{X} is computed by

$$\mathbf{X}^{ij} = \frac{1}{2} (\mathbf{X}^i + \mathbf{X}^j) \quad (10)$$

$$\mathbf{X}^i = \mathbf{t}_i + s_i^{ij} \mathbf{m}_i^{\mathbf{X}}, \quad \mathbf{X}^j = \mathbf{t}_j + s_j^{ij} \mathbf{m}_j^{\mathbf{X}},$$

with $\mathbf{t}_{i,j}$ being the projection centers of image i, j , $\mathbf{m}_{i,j}^{\mathbf{X}}$ being the unit-length vector from $\mathbf{t}_{i,j}$ to the respective image coordinates of point \mathbf{X} in image i, j and $s_{i,j}^{ij}$ being scale factors for image i, j that define the distance from $\mathbf{t}_{i,j}$ to $\mathbf{X}^{i,j}$, respectively (see Fig. 2 top left). $s_{i,j}^{ij}$ can be computed in a linear adjustment, finding the minimal distance between two skew rays given by $\mathbf{m}_{i,j}^{\mathbf{X}}$. These rays can also be described using the relative translation direction \mathbf{t}_{ij}^0 . All three rays are known, hence we can compute the rotation matrices rotating \mathbf{t}_{ij}^0 to $\mathbf{m}_{i,j}^{\mathbf{X}}$, i.e. $\mathbf{m}_i^{\mathbf{X}} = \mathbf{R}_{\alpha_i^{ij}} \mathbf{t}_{ij}^0$ and $\mathbf{m}_j^{\mathbf{X}} = \mathbf{R}_{\alpha_j^{ij}} (-\mathbf{t}_{ij}^0)$ to derive the following equation:

$$\mathbf{X}^{ij} = \frac{1}{2} (\mathbf{t}_i + s_i^{ij} \mathbf{R}_{\alpha_i^{ij}} \mathbf{t}_{ij}^0 + \mathbf{t}_j + s_j^{ij} \mathbf{R}_{\alpha_j^{ij}} (-\mathbf{t}_{ij}^0)) \quad (11)$$

Assuming point \mathbf{X} is also observed in image k (and a relative orientation $\{\mathbf{R}_{jk}^0, \mathbf{t}_{jk}^0\}$ has been estimated between j and k) there is another solution for \mathbf{X} given by:

$$\mathbf{X}^{jk} = \frac{1}{2} (\mathbf{t}_j + s_j^{jk} \mathbf{R}_{\alpha_j^{jk}} \mathbf{t}_{jk}^0 + \mathbf{t}_k + s_k^{jk} \mathbf{R}_{\alpha_k^{jk}} (-\mathbf{t}_{jk}^0)) \quad (12)$$

With $\mathbf{t}_{ij}^0 = \mathbf{t}_j - \mathbf{t}_i$ and analogously for all relative translations we equate Eq. (11) with (12) and derive

$$\mathbf{t}_i + \mathbf{C}_i (\mathbf{t}_j - \mathbf{t}_i) + \mathbf{t}_j + \mathbf{C}_j^{ij} (\mathbf{t}_i - \mathbf{t}_j) = \mathbf{t}_j + \mathbf{C}_j^{jk} (\mathbf{t}_k - \mathbf{t}_j) + \mathbf{t}_k + \mathbf{C}_k (\mathbf{t}_j - \mathbf{t}_k) \quad (13)$$

$$\Leftrightarrow (\mathbf{C}_i - \mathbf{C}_j^{ij}) (\mathbf{t}_j - \mathbf{t}_i) + \mathbf{t}_i + \mathbf{t}_j = (\mathbf{C}_j^{jk} - \mathbf{C}_k) (\mathbf{t}_k - \mathbf{t}_j) + \mathbf{t}_j + \mathbf{t}_k, \quad (14)$$

with

$$\mathbf{C}_i = s_i^{ij} \mathbf{R}_{\alpha_i^{ij}}, \quad \mathbf{C}_j^{ij} = s_j^{ij} \mathbf{R}_{\alpha_j^{ij}},$$

$$\mathbf{C}_j^{jk} = s_j^{jk} \mathbf{R}_{\alpha_j^{jk}}, \quad \mathbf{C}_k = s_k^{jk} \mathbf{R}_{\alpha_k^{jk}}.$$

From Eq. (14) we build the linear homogeneous constraint:

$$\mathbf{t}_i (\mathbf{C}_j^{ij} - \mathbf{C}_i + \mathbf{I}_{3 \times 3}) + \mathbf{t}_j (\mathbf{C}_i - \mathbf{C}_j^{ij} + \mathbf{C}_j^{jk} - \mathbf{C}_k) \dots + \mathbf{t}_k (\mathbf{C}_k - \mathbf{C}_j^{jk} - \mathbf{I}_{3 \times 3}) = \mathbf{0}. \quad (15)$$

Eq. (15) relates to the pairwise combination ij and jk . For other combinations (like $\{ij, ik\}$ or $\{ik, jk\}$) the constraint has to be adapted. If the point is measured in more than three images a constraint can be formulated for every independent triplet of images. However, most likely there will not be a single point that is measured in all images in V . Thus, in order to be able

¹In the local coordinate system mentioned above this would mean that $\mathbf{t}_i = [0, 0, 0]^T$ and $\mathbf{t}_j = \mathbf{R}_j^{*+} \mathbf{t}_{ij}^0$

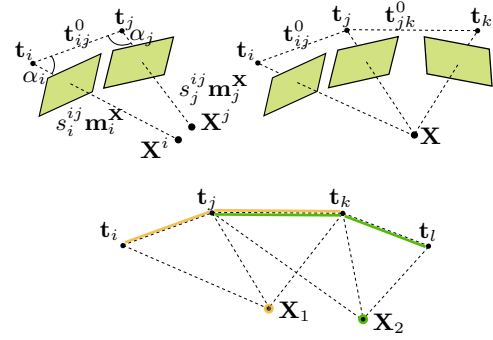


Figure 2. Geometry of spatial intersection minimizing the distance between two skew lines $\mathbf{m}_{i,j}^{\mathbf{X}}$ (upper left), constraint for an image triplet and two individual spatial intersections of \mathbf{X} to meet at the same position (upper right) and overlapping point tracks allow the estimation of translations which are not connected with a single point (bottom).

to estimate an absolute translation for every image, overlapping point tracks have to be found (see Fig. 2 bottom).

All constraints that are composed as described in Eq. (15) are stacked into a single matrix \mathbf{C} in order to solve the following linear optimization problem

$$\begin{aligned} & \text{minimize} && \|\mathbf{C}\mathbf{t}\| \\ & \{\mathbf{t}_1, \dots, \mathbf{t}_n\} && \\ & \text{subject to} && \|\mathbf{t}\| = 1, \end{aligned} \quad (16)$$

with \mathbf{t} being a $[3n \times 1]$ vector including all unknown translations. The norm-constraint in (16) avoids the obvious solution $\mathbf{t} = \mathbf{0}$ and defines the scale of the solution which can be selected arbitrarily. Problem (16) minimizes a geometric error because the linear constraints concern the distance between the intersection rays. The solution to (16) is computed with a SVD of \mathbf{C} , more precisely taking the rightmost column of \mathbf{V} , of course excluding zero-singular values.

5.2 Outlier Elimination from Spatial Intersection

The solution to (16) is the least squares solution of a homogeneous linear equation system. It is sensitive to outliers which may occur due to matching errors or errors in the relative orientations used. Cui et al. (2015) perform a L_1 optimization which clearly improves robustness but, nevertheless, outliers remain in the set of constraints. Since we perform an exhaustive filtering of relative rotations for rotation averaging (see Sec. 3.2) and transfer this result to relative translations we argue that these translations are outlier-free. Note that we already eliminate outliers in the homologous points during the estimation of relative orientations (see Sec. 3.1). Before translation estimation, however, rotations have been averaged (see Sec. 4) and points that withstood elimination in the first place are not necessarily inliers anymore. Thus, our goal is to identify and eliminate outliers in the homologous points before (16) is solved. A simple idea to achieve this goal is to consider the reprojection error of each individual local spatial intersection and exclude those observations whose reprojection error exceeds a certain threshold τ_r . The reprojection error is high if either the matching or the relative orientation (or both) is wrong. Because we precluded the latter to be correct, only the former option is possible. Note that this means that we are also able to detect those false relative orientations which would be apparent if all points for one image pair fail our reprojection constraint. However, this did not happen in any of our experiments.

Although this filtering works fine in most occasions, two critical aspects have to be taken care of. First, it may happen that one image does not have any constraints at all. In this case the affected image has to be deleted from V . Second, it is necessary that all images are connected as depicted in Fig. 2 bottom, i.e. single point tracks have to overlap in the edges of the graph. This has to be checked after the filtering. In case parts of the graph are not connected sufficiently, the estimation can be split into subproblems and the individual solutions can be merged afterwards, e.g. using structure information, provided that all vertices are covered.

6 EXPERIMENTS

In this section we present experiments we conducted with our model. First, we examine the robustness of our approach based on a synthetic set of images (Sec. 6.1). Thereafter in Sec. 6.2, we evaluate different versions of our model by means of the well-known benchmark datasets (Strecha et al., 2008) and compare our results to recent related works.

6.1 Robustness

In a first experiment we evaluate the robustness of our model. We use a linear image sequence in which 50 images are simulated along a 1-dimensional path, all pointing in the same direction. In total 2000 object points are generated, distributed as depicted in Fig. 3. The projections into the images are corrupted with an artificial normally distributed noise of 0.5px. It is important to note that the density of the resulting graph (i.e. the relation between the existing and the maximal possible number of edges) is approximately 0.34, because the ability to detect outliers in the graph depends on the redundancy of relative orientations. In this setting we compare different rates of outliers in the relative rotations. An outlier is generated by multiplying the estimated relative rotation with a random rotation matrix with $15^\circ \leq \{\omega, \phi, \kappa\} \leq 345^\circ$. For every investigated outlier rate, 100 individual iterations were conducted in which the set of outliers was randomly generated based on the corresponding outlier rate. We select $\tau_s = 5^\circ$ and $\tau_c = 1.5$.

As a state-of-the-art approach we select the LIRA and L1-IRLS model of Chatterjee and Govindu (2013)² and compare our results to theirs. The evaluation of LIRA and L1-IRLS is based on the same relative rotations we use for our model. Results are depicted in Fig. 4, note the logarithmic scale of the y-axis. Outlier configurations which lead to non-estimable rotations are not considered for all compared models. Until 40% of outliers in the relative rotations our model produces accurate results which implies that all defective rotations are filtered out. At 45% there is one case in which not all outliers have been removed which affects the mean angular error significantly. However, the mean angular error is constantly smaller than in both compared approaches for all investigated rates of outliers and also produces acceptable results at highest outlier rates whose angular error is in the range of a few $^\circ$. The false positive rate is around 10% at 50% outliers.

6.2 Benchmark Datasets

In this section we compare different settings of our model based on benchmark data with known ground truth (Strecha et al., 2008). The interior orientation is taken from EXIF

²This efficient model was used to estimate global rotations by several recent publications (Wilson and Snavely, 2014; Özyesil and Singer, 2015; Cui et al., 2015). The source code is available online

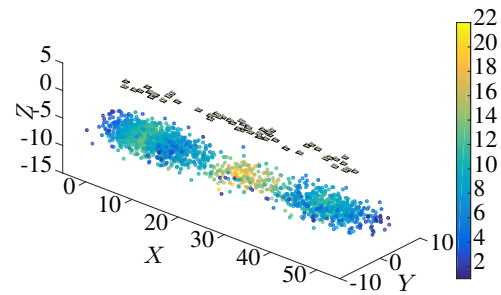


Figure 3. Synthetic image sequence, object points are colored by the number of images they are observed in.

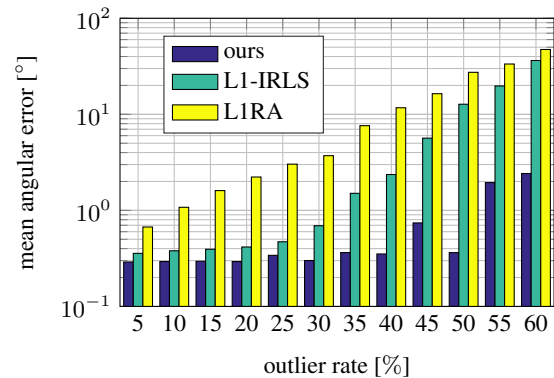


Figure 4. Comparison of our results (blue) with results from Chatterjee and Govindu (2013) (green and yellow) showing the mean angular error to the ground truth as a function of the rate of outliers in the relative rotations.

information³. In Tab. 1 we listed the mean angular error of our rotation estimation after graph optimization (GO), $SO(3)$ -relaxed SDP (SDP), least squares optimization (ls) and using unit weights instead of covariance information (I). Again we compared our results with the L1-IRLS model of Chatterjee and Govindu (2013) and three further approaches from related work. One can see that the local optimization has a strong effect on the accuracy of the estimated rotations whereas the usage of covariance-based weights only sometimes has a marginal positive effect on the results, for *castle-P30* there is a small negative effect. We argue the positive effect would be visible for more complex image sets in which the quality of the relative rotations is more diverse. Moreover, there is no gain between GO+ls and GO+SDP+ls except for *Herz-Jesu-P25*. As in the synthetic example in Sec. 6.1 our model achieves more accurate results than (Chatterjee and Govindu, 2013). This effect is almost negligible for *fountain-P11* and *Herz-Jesu-P25*, however the two castle-datasets show a significant improvement. Probably this is due to the higher amount of outliers in the relative orientations, caused by repetitive structure. We also compared our results to the rotations estimated in Jiang et al. (2013), Jiang et al. (2015) and Arie-Nachimson et al. (2012) where our model achieves superior results before and especially after bundle adjustment. However, their results are taken from the respective papers hence the shown values are based on a different set of relative orientations.

An evaluation of the global translations is given in Tab. 2. In all our experiments we set the maximum number of point tracks to 30 while one track includes a minimal set of points so that

³As initial values for the principal point we used the center of the image and the focal length from EXIF meta-information

	before bundle adjustment							after bundle adjustment		
	GO-SDP	GO-ls	GO-SDP-ls	GO-SDP-ls-I	L1-IRLS	(1)	(2)	GO-SDP-ls	(1)	(3)
<i>fountain-P11</i>	0.282	0.249	0.249	0.250	0.289	0.517	0.45	0.027	0.195	0.420
<i>Herz-Jesu-P25</i>	0.480	0.207	0.206	0.206	0.270	0.573	0.39	0.021	0.188	0.348
<i>castle-P19</i>	1.363	0.647	0.647	0.676	1.348	-	-	0.076	-	-
<i>castle-P30</i>	1.200	0.583	0.583	0.577	0.724	1.651	0.96	0.039	0.480	-

Table 1. Mean angular error of different versions of our model: graph optimization (GO), local least squares adjustment (ls), $SO(3)$ -relaxed SDP (SDP), unit weights (I). We compared our results to the robust L1-IRLS model of Chatterjee and Govindu (2013), the global method in Jiang et al. (2013) (1), the DSE(BA₃) in Jiang et al. (2015) (2) and the spectral method in Arie-Nachimson et al. (2012) (3). Note that no results for *castle-P19* are provided in (1) and (2) and additionally no results for *castle-P30* in (3). The rightmost column shows the result after bundle adjustment. All errors are given in [°].

	before bundle adjustment				after bundle adjustment				
	TE	TE-SI	(1)	(2)	TE-SI	(1)	(2)	(3)	(4)
<i>fountain-P11</i>	0.038	0.035	0.053	0.072	0.007	0.014	0.011	0.007	0.027
<i>Herz-Jesu-P25</i>	0.109	0.083	0.106	0.061	0.013	0.064	0.056	0.026	0.052
<i>castle-P19</i>	4.873	0.428	-	-	0.081	-	-	-	-
<i>castle-P30</i>	4.688	1.312	1.158	1.620	0.044	0.235	0.200	0.1667	-

Table 2. Mean translation error of different versions of our model given in [m]: translation estimation (TE), spatial intersection outlier elimination (SI). We compared our results to the results of the global method in Jiang et al. (2013) (1) and the DSE(BA₃) in Jiang et al. (2015) (2). The five rightmost columns show the results after bundle adjustment and an additional comparison to Cui et al. (2015) (3) and Arie-Nachimson et al. (2012) (4).

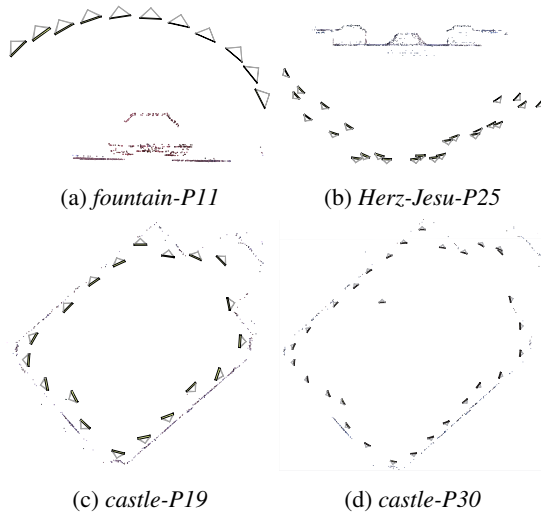


Figure 5. Visualization of the orientation results of our model with BA on benchmark image sequences.

every image is connected. We compare two different versions of our model, with and without using outlier elimination by spatial intersection (SI). We set the threshold for the outlier elimination to $\tau_r = 2\text{px}$. It can be seen that an outlier elimination in the homologous points has a positive effect especially on the more complex datasets *Herz-Jesu-P25*, *castle-P19* and *castle-P30*. The real benefit of our outlier elimination reveals itself after bundle adjustment in comparison with the results of Cui et al. (2015) (3). While both models performs equally well on *fountain-P11*, our model outperforms Cui et al. (2015) for *Herz-Jesu-P25* and *castle-P30*. A visualization of the estimated poses after bundle adjustment can be seen in Fig. 5.

7 CONCLUSION

In this paper we presented a novel approach for a convex estimation of global image orientations from relative orientations. Our model is divided into three essential steps, the estimation

of a consistent set of relative orientations, the estimation of global rotation parameters and finally the estimation of global translations. The main contributions in this paper are a novel heuristic approach to filter relative rotations and a robustification of a linear estimation of global translation parameters based on local spatial intersection. Our experiments show that the filtering of relative rotations works well on synthetic and real data and that our approach outperforms a state-of-the-art robust rotation averaging algorithm. The simple way in which we eliminate remaining outliers in the homologous points achieves a significant benefit as we showed with real benchmark data. The combination of a $SO(3)$ -relaxed SDP with a subsequent local optimization shows promising results. We think the improvement will be more significant for larger data sets.

In the future we will focus on processing unordered datasets from the Internet and in the course of this try to find effective approaches to formulate criteria for the generation of an appropriate MST in order to cover the increasing computational effort which greatly affects the orientation quality. We will also conduct further experiments on the benchmark data using ground truth interior orientation to be able to show additional comparisons to various related approaches.

APPENDIX

Relation between problem (3) and (6)

Let $d\mathbf{R}_{ij} = \mathbf{R}_i^T \mathbf{R}_j - \mathbf{R}_{ij}$. Then:

$$\begin{aligned} \|d\mathbf{R}_{ij}\|_F &= \sqrt{\text{tr}(d\mathbf{R}_{ij} d\mathbf{R}_{ij}^T)} \\ &= \sqrt{\text{tr}((\mathbf{R}_i^T \mathbf{R}_j - \mathbf{R}_{ij})(\mathbf{R}_j^T \mathbf{R}_i - \mathbf{R}_{ij}^T))} \\ &= \sqrt{\text{tr}(2\mathbf{I} - 2\mathbf{R}_i^T \mathbf{R}_j \mathbf{R}_{ij}^T)}, \end{aligned} \quad (17)$$

$$\text{because } \text{tr}(\mathbf{R}_i^T \mathbf{R}_j \mathbf{R}_{ij}^T) = \text{tr}(\mathbf{R}_{ij} \mathbf{R}_j^T \mathbf{R}_i).$$

We can extract the constant part from (17) and square it to derive the final form as in (6): $-\text{tr}(\mathbf{R}_i^T \mathbf{R}_j \mathbf{R}_{ij}^T) = -\text{tr}(\mathbf{M}_{ij} \mathbf{R}_{ij}^T)$.

REFERENCES

- Agarwal, S., Snavely, N., Simon, I., Seitz, S. and Szeliski, R., 2009. Building rome in a day. In: Proc. IEEE International Conf. on Computer Vision (ICCV), pp. 72–79.
- Arie-Nachimson, M., Kovalsky, S. Z., Kemelmacher-Shlizerman, I., Singer, A. and Basri, R., 2012. Global motion estimation from point matches. In: Proc. of the International Conf. on 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), IEEE, pp. 81–88.
- Arrigoni, F., Rossi, B. and Fusiello, A., 2015. Robust and efficient camera motion synchronization via matrix decomposition. In: Proc. International Conf. on Image Analysis and Processing (ICIAP), Springer, pp. 444–455.
- Chatterjee, A. and Govindu, V. M., 2013. Efficient and robust large-scale rotation averaging. In: Proc. IEEE International Conf. on Computer Vision (ICCV), pp. 521–528.
- Cui, Z., Jiang, N., Tang, C. and Tan, P., 2015. Linear global translation estimation with feature tracks. In: Proc. British Machine Vision Conference (BMVC).
- Enqvist, O., Kahl, F. and Olsson, C., 2011. Non-sequential structure from motion. In: Proc. IEEE International Conference on Computer Vision Workshops (ICCV Workshops), IEEE, pp. 264–271.
- Fischler, M. A. and Bolles, R. C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24(6), pp. 381–395.
- Förstner, W., Wrobel, B., Paderes, F., Craig, R., Fraser, C. and Dolloff, J., 2004. Manual of Photogrammetry. American Society for Photogrammetry and Remote Sensing, chapter 11, Analytical photogrammetric operations, pp. 763–948.
- Fredriksson, J. and Olsson, C., 2012. Simultaneous multiple rotation averaging using lagrangian duality. In: Asian Conference on Computer Vision (ACCV), Springer, pp. 245–258.
- Gilmore, R., 1974. Baker-campbell-hausdorff formulas. *Journal of Mathematical Physics* 15(12), pp. 2090–2092.
- Govindu, V. M., 2001. Combining two-view constraints for motion estimation. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 2, IEEE, pp. II–218.
- Govindu, V. M., 2004. Lie-algebraic averaging for globally consistent motion estimation. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 1, IEEE, pp. I–684.
- Hartley, R., Aftab, K. and Trumpp, J., 2011. L1 rotation averaging using the Weiszfeld algorithm. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 1, pp. 3041–3048.
- Hartley, R. and Schaffalitzky, F., 2004. L_∞ minimization in geometric reconstruction problems. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 1, pp. I–504–I–509 Vol.1.
- Hartley, R., Trumpp, J., Dai, Y. and Li, H., 2013. Rotation averaging. *International Journal of Computer Vision (IJCV)* 103(3), pp. 267–305.
- Horowitz, M. B., Matni, N. and Burdick, J. W., 2014. Convex relaxations of $se(2)$ and $se(3)$ for visual pose estimation. In: Proc. IEEE International Conference on Robotics and Automation (ICRA), Vol. abs/1401.3700.
- Jiang, N., Cui, Z. and Tan, P., 2013. A global linear method for camera pose registration. In: Proc. IEEE International Conf. on Computer Vision (ICCV), IEEE, pp. 481–488.
- Jiang, N., Lin, W.-Y., Do, M. N. and Lu, J., 2015. Direct structure estimation for 3d reconstruction. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2655–2663.
- Kahl, F. and Hartley, R., 2008. Multiple-view geometry under the L_∞ -norm. *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)* 30(9), pp. 1603–1617.
- Kanatani, K., 1990. Group-theoretical methods in image understanding. Vol. 2, Springer.
- Ke, Q. and Kanade, T., 2007. Quasiconvex optimization for robust geometric reconstruction. *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)* 29(10), pp. 1834–1847.
- Martinec, D. and Pajdla, T., 2007. Robust rotation and translation estimation in multiview reconstruction. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, pp. 1–8.
- Moulon, P., Monasse, P., Marlet, R. et al., 2013. Global fusion of relative motions for robust, accurate and scalable structure from motion. In: Proc. IEEE International Conf. on Computer Vision (ICCV).
- Olsson, C. and Kahl, F., 2010. Generalized convexity in multiple view geometry. *Journal of Mathematical Imaging and Vision* 38(1), pp. 35–51.
- Özyesil, O. and Singer, A., 2015. Robust camera location estimation by convex programming. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2674–2683.
- Özyesil, O., Singer, A. and Basri, R., 2015. Stable camera motion estimation using convex programming. *SIAM Journal on Imaging Sciences* 8(2), pp. 1220–1262.
- Pollefeys, M., Van Gool, L., Vergauwen, M., Verbiest, F., Cornelis, K., Tops, J. and Koch, R., 2004. Visual modeling with a hand-held camera. *International Journal of Computer Vision (IJCV)* 59(3), pp. 207–232.
- Saunders, J., Parrilo, P. and Willsky, A., 2014. Semidefinite relaxations for optimization problems over rotation matrices. In: Proc. IEEE Conference on Decision and Control (CDC), pp. 160–166.
- Sinha, S. N., Steedly, D. and Szeliski, R., 2012. A multi-stage linear approach to structure from motion. In: Trends and Topics in Computer Vision, Springer, pp. 267–281.
- Strecha, C., von Hansen, W., Van Gool, L., Fua, P. and Thoennessen, U., 2008. On benchmarking camera calibration and multi-view stereo for high resolution imagery. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, pp. 1–8.
- Weiszfeld, E. and Plastria, F., 2009. On the point for which the sum of the distances to n given points is minimum. *Annals of Operations Research* 167(1), pp. 7–41.
- Wilson, K. and Snavely, N., 2014. Robust global translations with 1dsfm. In: Proc. of the European Conf. on Computer Vision (ECCV), Springer, pp. 61–75.
- Wu, C., Agarwal, S., Curless, B. and Seitz, S. M., 2011. Multicore bundle adjustment. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, pp. 3057–3064.
- Zach, C., Klopschitz, M. and Pollefeys, M., 2010. Disambiguating visual relations using loop constraints. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1426–1433.