

# GLOBAL ROTATION ESTIMATION USING WEIGHTED ITERATIVE LIE ALGEBRAIC AVERAGING

M. Reich\*, C. Heipke

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

**KEY WORDS:** image orientation, pose estimation, rotation averaging, Lie algebra

## ABSTRACT:

In this paper we present an approach for a weighted rotation averaging to estimate absolute rotations from relative rotations between two images for a set of multiple overlapping images. The solution does not depend on initial values for the unknown parameters and is robust against outliers. Our approach is one part of a solution for a global image orientation. Often relative rotations are not free from outliers, thus we use the redundancy in available pairwise relative rotations and present a novel graph-based algorithm to detect and eliminate inconsistent rotations. The remaining relative rotations are input to a weighted least squares adjustment performed in the Lie algebra of the rotation manifold  $SO(3)$  to obtain absolute orientation parameters for each image. Weights are determined using the prior information we derived from the estimation of the relative rotations. Because we use the Lie algebra of  $SO(3)$  for averaging no subsequent adaptation of the results has to be performed but the lossless projection to the manifold. We evaluate our approach on synthetic and real data. Our approach often is able to detect and eliminate all outliers from the relative rotations even if very high outlier rates are present. We show that we improve the quality of the estimated absolute rotations by introducing individual weights for the relative rotations based on various indicators. In comparison with the state-of-the-art in recent publications to global image orientation we achieve best results in the examined datasets.

## 1. INTRODUCTION

In this work we tackle a sub-problem for the determination of three dimensional information from overlapping images or image sequences, i.e. the estimation of absolute rotations. The term 'absolute rotations' refers to rotations given for each image in a local coordinate system, whereas 'relative rotations' denote rotations from one image to a next overlapping counterpart. This is a basic task in photogrammetry and related sciences as it is an important prerequisite for many different applications. Many well known and highly performant works use sequential approaches to jointly solve for all the unknown orientation parameters (Agarwal et al., 2009), (Snavely et al., 2006). Initial values for a final nonlinear bundle adjustment are computed starting with a small subset of images (often only two or three). This set is incrementally enlarged by subsequent resection, spatial intersection and intermediate bundle adjustment in order to avoid the solution to suffer from drift and/or divergence. The quality of the initial values generated in this manner depends on two important factors: the initial subset of images and the ordering in which images are added to the existing block. These factors are determined based on heuristics that are rarely useful without further information about the image's position in object space like geographic meta-data.

In our work we present and analyse the first part of a non-sequential approach for the estimation of image orientation parameters that works with homologous points, i.e. tie points in overlapping images, and relative rotations between pairs of images and is independent on initial values for the unknowns. Following the common terminology we call this non-sequential approach a global approach (Arie-Nachimson et al., 2012), (Jiang et al., 2013), (Moulon et al., 2013). The workflow for the estimation of image orientation parameters is outlined in Fig. 1. In this work we focus on one relevant step, the estimation of absolute rotations. We as-

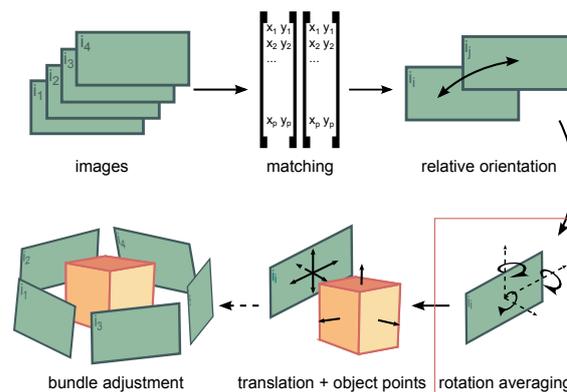


Figure 1: Workflow of our global image orientation approach. In this work we focus on the highlighted part - rotation averaging.

sume that homologous points have been found and relative rotations have been computed but both are not necessarily free from outliers. We developed an effective outlier detection algorithm for relative rotations based on a graph structure. In this graph absolute rotations are depicted by nodes that are connected by edges if a relative rotation between the two respective images has been computed. Normally, images have an overlap not only to one but to more neighbouring images. Thus, there will be more edges than nodes in the graph and each edge bears a constraint on the nodes it connects, i.e. providing redundant information regarding the estimation of absolute rotations. We use this redundancy in order to obtain a good initialisation for a linearised least squares adjustment and to detect and eliminate outliers in the relative rotations that may remain, e.g. after the estimation of the essential matrix or trifocal tensor. The least squares adjustment is performed via projection to the tangent space of the rotation Lie group  $SO(3)$  - the Lie algebra  $so(3)$  - and a linearisation in order to estimate the unknown absolute rotations. We integrate

\*Corresponding author

prior knowledge about the quality of relative rotations, incorporated as weights on individual relative rotations in the adjustment, to further enhance the performance of the rotation averaging.

The remainder of this paper is organised as follows: In the following section we give an overview of the related work and our contribution. Section 3 is dedicated to our approach. We focus on the description of the robust estimation of absolute rotations. For the estimation of translations and object points we refer to Reich and Heipke (2014). In Section 4 we will evaluate our approach based on synthetic and real data and finally give a conclusion of our work in Section 5.

## 2. RELATED WORK

The work on image orientation in recent years was highly supported by the growing number of publicly available image data and increasing computing power. Agarwal et al. (2009), Snavely et al. (2006), for instance, developed ways to compute image orientations of thousands of images from online image galleries. To keep the computational effort tractable they match images in a sequential manner starting with an initial sub-set of images in order to avoid having to match every possible combination of image pairs. The initial orientation parameters computed in this way, however, depend on the initial sub-set and the ordering in which images are added and may lead to a poor overall accuracy of the final image orientations. This is why global methods recently gained a lot of attention (Martinec and Pajdla, 2007), (Enqvist et al., 2011), (Arie-Nachimson et al., 2012), (Sinha et al., 2012), (Jiang et al., 2013), (Moulon et al., 2013). Often these approaches work in a two-step manner, first estimating absolute rotations and then translations and/or object points.

One of the first works exploring the idea of estimating absolute rotations from relative ones is given by Govindu (2001). The problem is solved using quaternions in a linear least squares estimation. The length-constraint stating that only unit quaternions represent valid rotations is neglected, however. A few years after his first work on rotation averaging Govindu proposed a new approach using Lie group theory estimating both, absolute rotations and translations from relative estimates in a non-linear optimisation (Govindu, 2004). Because errors during the estimation of essential matrices occur regularly there are works investigating more robust methods for rotation averaging like using the  $L_1$  norm (Dai et al., 2010), (Chatterjee and Govindu, 2013), (Hartley et al., 2011), (Hartley et al., 2013). Hartley et al. (2011) and Hartley et al. (2013), for instance, describe a sequential rotation averaging algorithm for a graph of absolute rotations in which one absolute rotation is averaged at a time and iteratively improved in the manner of a distributed consensus. They use the Weiszfeld algorithm (Weiszfeld, 1937)<sup>1</sup> that minimises the  $L_1$  norm. Furthermore, Hartley et al. (2013) give a detailed proof for convexity of averaging a single rotation and investigate many mathematical aspects about rotation averaging and convexity on manifolds

Instead of the  $L_1$  norm Reich and Heipke (2014) use a Huber cost function which shows similar robustness while being computationally less demanding. Many works also pursue the idea of eliminating outliers before the averaging, e.g. using minimum spanning trees (minimum sub-graphs in which every node can be reached from every other) (Govindu, 2006), using a Bayesian network and cycles in the view-graph (Zach et al., 2010), or starting from a most reliant spanning tree, and adding new edges iteratively (Enqvist et al., 2011). All of these works assume a certain

<sup>1</sup>For this work a translated version exists (Weiszfeld and Plastria, 2009)

maximum length of cycles or the correctness of specific relative rotations.

## 2.1 Contribution

In this work we present a new approach for the robust estimation of absolute rotations for a sequence of overlapping images. Our detection of outliers in the relative rotations does not suffer from cycle length limitation and yields promising results. We show how we improve the estimation of absolute rotations by introducing weights on relative rotations based on number, distribution and reprojection quality of homologous points. Our approach yields comparable and partly significantly better results with respect to state-of-the-art approaches.

## 3. ESTIMATING ABSOLUTE ROTATIONS

In this section we describe how we estimate an absolute rotation for each out of  $n$  images from redundant relative rotations existing for pairs of images. An absolute rotation  $\mathbf{R}_j$  can be propagated from  $\mathbf{R}_i$  by

$$\mathbf{R}_j = \mathbf{R}_i \mathbf{R}_{ij}, \quad \forall (i, j) \in N = 1, \dots, n; i \neq j, \quad (1)$$

using the relative rotation  $\mathbf{R}_{ij}$  between the two images  $i$  and  $j$ . Homologous points in overlapping images and relative rotations from essential matrices are assumed to be known and their computation is not focus of this work.

### 3.1 Preliminaries

Rotation matrices form an algebraic group, called the special orthogonal group  $SO(3)$ , a Lie group with multiplication as operator (for a comprehensive overview about algebraic groups in image processing the reader is referred to Kanatani (1990)). Every multiplication of two rotation matrices will result in another rotation matrix, the identity element is the identity matrix  $\mathbf{I}_{3 \times 3}$  and the inverse element is equal to the transposed rotation matrix  $\mathbf{R}^{-1} = \mathbf{R}^T$ .  $SO(3)$  consists of all matrices in  $\mathbb{R}^{3 \times 3}$  whose three columns span an orthonormal basis and whose determinant is equal to 1.

Rotation matrices also form a differentiable manifold which is inherent in every Lie group. Averaging on this manifold, i.e. finding the closest absolute rotation with respect to some redundant propagations, cannot be performed in the same way as in Euclidean space, because on the manifold only multiplication is defined; a summation of two rotation matrices leads to a result which is not a member of  $SO(3)$ . Hence, the computation of an ordinary arithmetic average of rotation matrices is prohibited. We use the fact that rotation matrices can be projected to their Euclidean tangent space, the Lie algebra  $\mathfrak{so}(3)$ , via the matrix logarithm  $\log(\cdot) : SO(3) \rightarrow \mathfrak{so}(3)$  (Hartley et al., 2013):

$$\log(\mathbf{R}) = [\mathbf{v}]_{\times}, \quad \mathbf{v} = \arcsin(\|\mathbf{w}\|_2) \frac{\mathbf{w}}{\|\mathbf{w}\|_2}, \quad (2)$$

$$\mathbf{w} = \frac{\mathbf{R} - \mathbf{R}^T}{2}.$$

$\mathfrak{so}(3)$  consists of all skew-symmetric matrices in  $\mathbb{R}^{3 \times 3}$

$$[\mathbf{v}]_{\times} = \begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix} \quad (3)$$

<sup>2</sup>We interpret every absolute rotation matrix  $\mathbf{R}$  as a rotation from image to object coordinate system

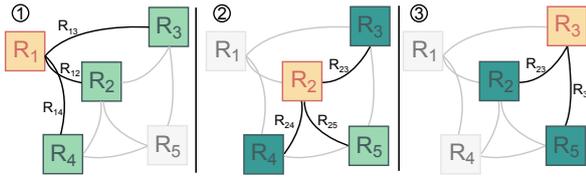


Figure 2: Construction of the graph and outlier detection: source for propagation (orange), rotations that are propagated to for the first time (light green), rotations that are propagated to and exhibit an estimation from previous propagation (dark green) and rotations that are not propagated to or stay fixed (grey).

and  $\mathbf{v}$  is equivalent to the three elements of the angle-axis representation of rotations  $\mathbf{v} = \alpha \tilde{\mathbf{v}}$  with angle  $\alpha$ ,  $\alpha \in \mathbb{R}$ , and rotation axis  $\tilde{\mathbf{v}} = (\tilde{v}_1, \tilde{v}_2, \tilde{v}_3)^T$ ,  $\|\tilde{\mathbf{v}}\| = 1$ . Thus, there exists a direct cast from matrix- to angle-axis representation with  $[\log(\cdot)]_{\times}^{-1} : SO(3) \rightarrow \mathbb{R}^3$ . A backprojection exists via the exponential map  $\exp[\cdot]_{\times} : \mathbb{R}^3 \rightarrow SO(3)$  using Rodrigues' rotation formula:

$$\exp([\mathbf{v}]_{\times}) = \mathbf{I} + \sin(\alpha) [\tilde{\mathbf{v}}]_{\times} + (1 - \cos(\alpha)) [\tilde{\mathbf{v}}]_{\times}^2 \quad (4)$$

### 3.2 Graph initialisation and outlier detection

The first part of our rotation estimation is the initialisation of the graph with absolute rotations being nodes, connected by edges if a relative rotation between the two respective images has been computed. In the beginning the nodes have not been assigned a value. During the initialisation of the graph we detect outliers in the relative rotations that are deleted before the second part, the iterative averaging. In order to differ outliers from inliers we use the geodesic metric that defines the angular similarity between rotations. This metric is defined as the shortest distance between two rotations on the manifold (the geodesic), projected into the tangent space:

$$d(\mathbf{R}_1, \mathbf{R}_2)_{geod} = d(\mathbf{R}_1^T \mathbf{R}_2, \mathbf{I})_{geod} = \left\| \left[ \log \mathbf{R}_1^T \mathbf{R}_2 \right]_{\times}^{-1} \right\|_2, \quad (5)$$

$\|\cdot\|_2$  is the Euclidean norm, defining the shortest distance in  $\mathbb{R}^3$ . Note that the identity matrix  $\mathbf{I}$  describes a rotation by  $\alpha = 0$ , hence, according to Eq. (2) its projection to tangent space results in the zero matrix  $\mathbf{0}_{3 \times 3}$ .

Starting from some node with a given rotation that stays fixed and defines a local coordinate system, we use Eq. (1) to propagate to all nodes in the graph. The procedure is outlined in Fig. 2. We refer to the node we are propagating from as the source node, the node we are propagating to is referred as the goal node. During propagation we differentiate two different situations for the goal node: A rotation has either been propagated before, or not. In the latter case propagation is straightforward, the former case implies that a cycle in the graph exists, involving the source- and the goal node. In this case we first compute the similarity of the two propagations using Eq. (5). If the similarity is high, i.e. the distance is below a certain threshold  $\theta$ ,  $d_{geod} \leq \theta$ , we compute the average of the two rotations minimising Eq. (5) with respect to the  $L_2$ -norm with the algorithm described in Hartley et al. (2013) (algorithm1, p. 16). All relative rotations inherent in the loop are subsequently classified as inliers. In case we do not find a high similarity, generally all relative rotations in the loop are potential outliers. In order to decrease the number of possible outliers we distinguish three different situations following the assumption that relative rotations are estimated independently:

1. both involved nodes (source and goal) have been averaged by at least two relative rotations: the relative rotation between the source- and the goal node causing the high distance  $d_{geod} > \theta$  is classified as outlier and rejected.
2. either the source- or the goal node has been averaged by at least two relative rotations: all relative rotations of the loop involved in the averaging of the respective node are rejected from the set of potential outliers.
3. a certain amount  $k$  of all propagations from the source node violate the similarity constraint  $d_{geod} \leq \theta$ : the source node obtains a wrong value and the relative rotation used to propagate this value to the source node is classified as outlier and is rejected. Consequently, the source node does not obtain a value anymore and has to be propagated again from another node.

Otherwise, the set of possible outliers is not decreased. In this case or in case 2, in which no final decision can be made, the goal node is affected by outliers in the relative rotations. The estimation is deferred until another propagation to this particular node is possible. Then we compare all available propagations and select consistent ones which allows for a subsequent classification of inliers and outliers. Note that we propagate only once from every node in the graph except case 3 is reached. In this case, as soon as another value is propagated to this particular node, propagation starts again. In this way we guarantee that the algorithm will converge with every node exhibiting a value, except the case one node is connected to other nodes by inconsistent edges only. Then the particular node is rejected from the subsequent least squares adjustment.

### 3.3 Iterative least squares adjustment

The second part of our rotation averaging approach is the adjustment of the relative rotations and the estimation of absolute rotations. For our iterative least squares adjustment we assume a graph that is free from any outlier in the relative rotations. Our algorithm follows in spirit the one of Govindu (2004) but we only estimate absolute rotations. We set the functional model according to Eq. (1) with  $\Delta \mathbf{R}_{ij}$  being the residual rotation matrix

$$\mathbf{R}_{ij} = \mathbf{R}_i^T \Delta \mathbf{R}_{ij} \mathbf{R}_j, \quad \forall (i, j) \in N = 1, \dots, n; i \neq j. \quad (6)$$

This model is non-linear in the unknown absolute rotations  $\mathbf{R}_i$  and  $\mathbf{R}_j$ . In order to linearise our functional model we use the projection to tangent space at the current estimation of absolute rotations:

$$\log(\mathbf{R}_{ij}) = \log\left(\mathbf{R}_i^T \Delta \mathbf{R}_{ij} \mathbf{R}_j\right). \quad (7)$$

Eq. (7) can be linearised using the first order approximation of the Baker-Campbell-Hausdorff-formula (BCH, see e.g. (Govindu, 2004)):  $\log(\mathbf{X}\mathbf{Y}) \approx \log(\mathbf{X}) + \log(\mathbf{Y})$ ,  $\mathbf{X}, \mathbf{Y} \in SO(3)$ . Then, Eq. (7) can be written as:

$$\begin{aligned} \Delta \mathbf{r}_{ij} &= \mathbf{r}_i - \mathbf{r}_j + \mathbf{r}_{ij}, \\ \Delta \mathbf{r}_{ij} &= \mathbf{r}_i - \mathbf{r}_j + \mathbf{r}_{ij}, \end{aligned} \quad (8)$$

with  $\mathbf{r} = \log(\mathbf{R})$ ,  $\mathbf{r} \in \mathfrak{so}(3)$  and  $\mathbf{r} = [\mathbf{v}]_{\times}^{-1}$ ,  $\mathbf{r} \in \mathbb{R}^3$ . This linearisation comes at a price. It is obvious that the higher order terms of the BCH-formula are not zero, because  $\mathbf{X}\mathbf{Y} - \mathbf{Y}\mathbf{X} \neq 0$ . Furthermore, the approximation is not bi-invariant, i.e.  $\log(\mathbf{S}\mathbf{X}) + \log(\mathbf{S}\mathbf{Y}) \neq \log(\mathbf{X}) + \log(\mathbf{Y})$  (Hartley et al., 2013). However, we made good experience with this approach and prefer it against other approaches.

The functional model in Eq. (8) is linear so that the design matrix can be easily constructed from positive and negative unit matrices  $\mathbf{I}_{3 \times 3}$ ,  $\mathbf{A} = [\mathbf{I}, -\mathbf{I}, \dots; \mathbf{I} \dots - \mathbf{I} \dots]$ . We use the reduced observations  $\mathbf{r}_{ij}^0 = [\log(\hat{\mathbf{R}}_i \mathbf{R}_{ij} \hat{\mathbf{R}}_j^T)]_x^{-1}$  that are computed using the current estimation of  $\mathbf{R}_i$  and  $\mathbf{R}_j$ ,  $\hat{\mathbf{R}}_i$  and  $\hat{\mathbf{R}}_j$ , respectively. By stacking the reduced observations into a vector we iteratively estimate improvements for the unknowns. The reduced observations are updated after every iteration. In order to take care of the rank deficiency we keep one absolute rotation fixed and exclude it from the adjustment. Generally, this is the rotation we fixed during graph initialisation (Section 3.2).

In general, we can derive information about the quality of relative rotations if they are computed from homologous points in overlapping image pairs or triplets. Thus, we use this information as a prior to perform a weighting on the relative rotations in order to improve the estimation. There are three possible indicators that can serve as a source for prior information:

1. number of homologous points used for the estimation of the fundamental- or essential matrix  $\tau_n$ .
2. distribution of the homologous points within image space. In order to describe the distribution we use the variance of the point coordinates with respect to their centre of gravity  $\tau_d$ .
3. reprojection error in image space using the fundamental- or essential matrix  $\tau_r$ .

Of course, each criterion itself might not be a self-sustaining indicator for the quality of the relative rotation, e.g. a small reprojection error does not coincide with well distributed or numerous homologous points. However, in our examples we use a minimum number of 40 homologous points in order to avoid minimal solutions with zero reprojection error.

The weighting matrix is constructed individually for every relative rotation in the manner described in Krarup et al. (1980), Klein and Förstner (1984).

$$\mathbf{P}_{ij} = \frac{1}{1 + (\lambda|v_{ij}|)^\nu} \mathbf{I}_{3 \times 3} \quad (9)$$

$$\lambda = \frac{1}{hw}$$

$$\nu = 4$$

Depending on the indicator one wants to use for the weighting, the reciprocal weights  $v_{ij}$  are computed:

$$v_{ij,n} = \frac{1}{\tau_n}, \quad v_{ij,d} = \frac{1}{\tau_d}, \quad v_{ij,r} = \tau_r \quad (10)$$

$hw$  is the halfweight of the weighting function, indicating the value of  $v_{ij}$  at which the relative rotations are assigned a weight of 0.5. This value also gives information about the steepness of the weighting function and has to be adjusted with respect to the reciprocal weights, because if the value is set too small, the system might become close to singular. On the other hand if  $hw$  is set too high the effect of weighting will only be marginal. We made good experience with setting  $hw$  according to the maximal reciprocal weight,  $hw = \frac{\max(v_{ij})}{2}$ . Note that we assume that outliers have been removed in the previous step (Sec. 3.2). Therefore, the weighting is not used to decide between good and bad relative rotations but only to introduce prior knowledge on the quality of the relative rotations. Matrix  $\mathbf{P}$  is set only once

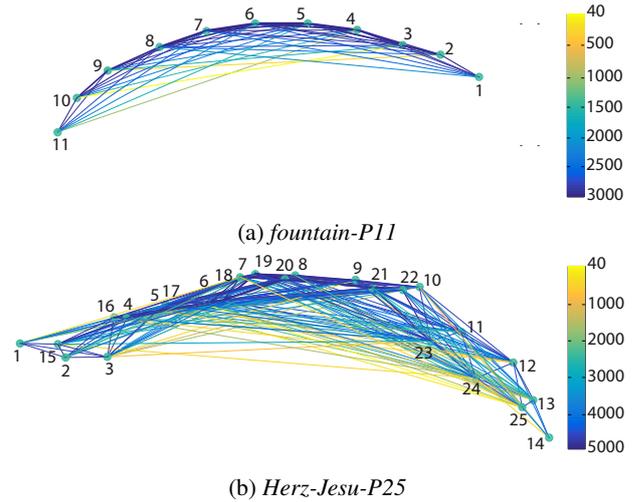


Figure 3: Visualisation of the graph. Colour and thickness of the edges refer to the number of homologous points found between each corresponding pair.

from all individual  $\mathbf{P}_{ij}$  and is not changed during the iterative estimation. The characteristics for each of the three indicators regarding the improvement of the estimation are evaluated in Sec. 4.

#### 4. EVALUATION

We evaluate our approach based on synthetic and real image sequences. The synthetic sequences are described below for each respective experiment. The real image sequences are the *fountain-P11* and *Herz-Jesu-P25* datasets (Strecha et al., 2008), consisting of 11 and 25 images, respectively. Both sequences are arranged in a half circle with cameras pointing towards the centre showing a fountain on a wall and a portal entrance (see Fig. 7). These datasets come with a ground truth for the orientation parameters generated using laserscanning and marked control points enabling an evaluation of our approach also with respect to related approaches of recent publications. The images are corrected for radial distortion. We downscaled the imagery by a factor of approx. 4.25 to keep the computational effort low which leads to a ground sampling distance of approx. 1 – 2cm. We assume non-calibrated cameras and use the information from the EXIF header for the interior orientation parameters. The graph describing the relative rotations for both sequences can be seen in Fig. 3.

In a first experiment we evaluate the performance of our outlier detection during the estimation of absolute rotations. As mentioned in Section 3.2 we have to define a threshold that separates good from bad relative rotations using Eq. (5). We set this threshold to  $\theta = 0.1[rad]$  which we have experimentally found to work well. The critical amount  $k$  is fixed at 50%. For this experiment we use synthetic data because we can easily generate outliers in the relative rotations. Twelve images are arranged in a circle all pointing towards the centre in which 100 object points are equally distributed within a cube. The base to height ratio between each neighbouring pair of images is approximately 0.5 and every object point is observed in every image. We added noise to the image coordinate observations, normally distributed with  $\sigma = 1px$ . The graph indicating the relative rotations is fully connected leading to a total number of 66 relative rotations. These rotation matrices are computed from essential matrices and a constant interior orientation. Outliers are generated by a left multiplication of the respective relative rotation matrix with a random rotation

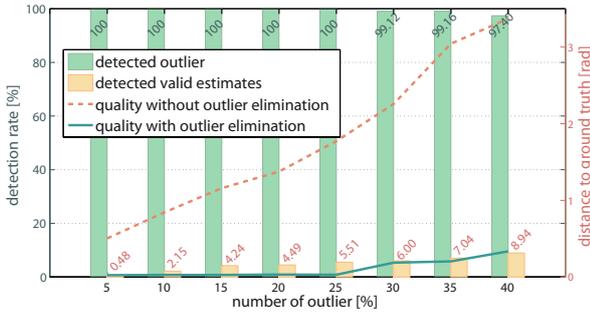


Figure 4: Outlier detection for various rates of outliers in the relative rotations, true positives (green bars) and false positives (orange bars). The curves show the accumulated distance of all estimated rotation matrices to their ground truth with and without outlier elimination.

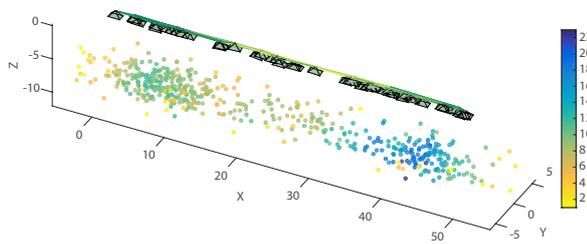
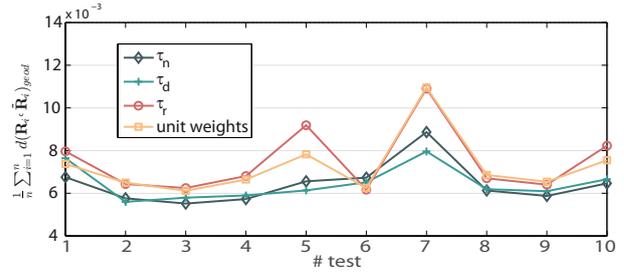


Figure 5: Synthetic arrangement of images used for the evaluation of weighting relative rotations.

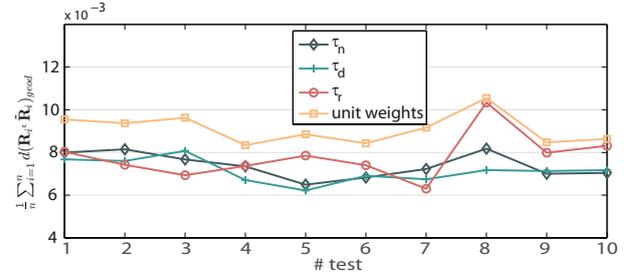
using Euler angles  $\{(\omega, \phi, \kappa) \in \mathbb{R} \mid 15^\circ \leq \omega, \phi, \kappa \leq 345^\circ\}$ . We vary the number of outliers between 5% and 40% of the relative rotations by a uniformly distributed sampling. For each number of outliers we repeated the detection and averaging 100 times using another independently generated configuration of outliers and take the average values. Results are depicted in Fig. 4. One can see that with up to 25% outliers in the relative rotations all outliers could be found whereas less than 6% of correct relative rotations are classified as outlier. With 40% outliers still a high number of them could be classified correctly. This is also visible in the two curves showing the accumulated distance of the estimated rotations to the ground truth. Note that in this experiment no weights were used because the symmetric configuration of images and object points does not lead to a measurable variation within the three indicators.

In our second experiment we compare the three different indicators, discussed in Section 3.3. We evaluate the rotation averaging results based on synthetic but also on the real data. As synthetic data we use a different configuration of images than in the first experiment to achieve a higher and more realistic variation in the three indicators for pairs of images. Now the sequence describes a rather linear movement of the camera in  $X$ -direction with 50 images, all pointing in a similar direction, e.g. simulating a single strip from airborne photogrammetry or a sequence from a mobile mapping device. We generated 500 object points forming two clusters with approx. 210 points each. The rest of the points is uniformly distributed over the whole scene. This arrangement is visualised in Fig. 5. The color of the object points refers to the number of images they are measured in.

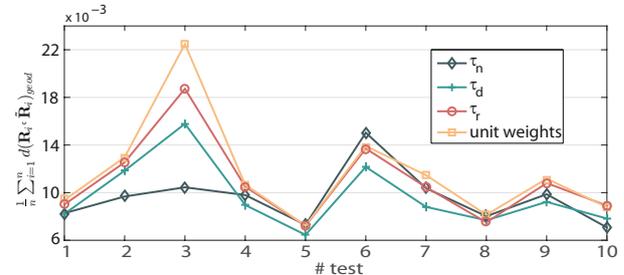
For every dataset we repeated the estimation 10 times and in every stage computed a new set of relative rotations using RANSAC algorithm (Fischler and Bolles, 1981). We always used the rotation with most connections in the graph as starting node for the propagation through the graph. This propagation serves as ini-



(a) Synthetic data



(b) fountain-P11



(c) Herz-Jesu-P25

Figure 6: Estimation results depicted as the average distance to the respective ground truth information  $\tilde{\mathbf{R}}$  in [rad] using different indicators responsible for weighting relative rotations and unit weights as comparison.

tialisation for the iterative least squares adjustment in which we perform 4 different estimations, three of them using one of the reciprocal weights defined in Eq. (10) each and one using constant unit weights. We compared the results to the ground truth information. The results are depicted in Fig. 6. Apparently, the results highly depend on the quality of the relative rotations which reflects itself in the variation of the results over the the ten repetitions, especially in the *Herz-Jesu-P25* sequence. When we compare the results of the individual weightings, clearly using one of the three indicators outperforms a unit weighting. This can be also manifested numerically (Tab. 1). Using the number of homologous points ( $\tau_n$ ) or the distribution of the points in image space ( $\tau_d$ ) as weights leads to the largest improvement with only marginal differences between each other. A weighting based on the reprojection error ( $\tau_r$ ) in the synthetic image sequence leads to a degradation of the results compared to unit weights. The improvement of using  $\tau_n$  or  $\tau_d$  for the weighting can also be seen in terms of a higher repeatability. The standard deviation of the results is significantly smaller than for unit weights, for example, and hence the dependency on the quality of the relative rotations is less.

In a third experiment we compare the results of our approach on the real datasets with respect to the ground truth and the results of similar approaches from recent publications. The orientation results of our approach are depicted in Fig. 7. The translations and

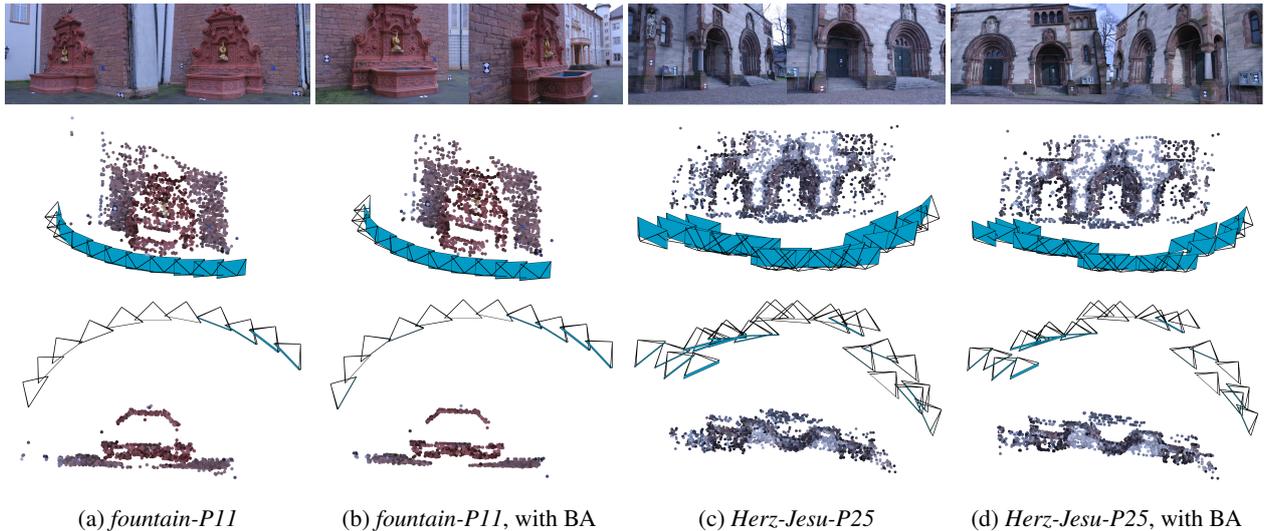


Figure 7: Visualisation of the orientation results of our approach on real image sequences, partly improved by a subsequent bundle adjustment (7b and 7d). The top row shows exemplary images.

|                      | $\tau_n$     | $\tau_d$     | $\tau_r$    | u. w.       |
|----------------------|--------------|--------------|-------------|-------------|
| synthetic data       | <b>0.369</b> | 0.370        | 0.430       | 0.416       |
| <i>fountain-P11</i>  | 0.424        | <b>0.409</b> | 0.447       | 0.521       |
| <i>Herz-Jesu-P25</i> | <b>0.549</b> | 0.556        | 0.627       | 0.666       |
| synthetic data       | $\pm 0.055$  | $\pm 0.045$  | $\pm 0.090$ | $\pm 0.081$ |
| <i>fountain-P11</i>  | $\pm 0.034$  | $\pm 0.031$  | $\pm 0.061$ | $\pm 0.040$ |
| <i>Herz-Jesu-P25</i> | $\pm 0.131$  | $\pm 0.159$  | $\pm 0.195$ | $\pm 0.250$ |

Table 1: Results of the rotation estimation using different weightings. Mean distance to ground truth (upper part) and standard deviation (lower part), both in  $[\circ]$ .

object points are estimated by the approach described in Reich and Heipke (2014) but using an iterative reweighted procedure in which we detect and delete outliers in the homologous points. Results in Fig. 7b and 7d are improved by a subsequent bundle adjustment. There is hardly any difference visible, the facades appear a bit more planar for the improved solution, visible in the top views in the bottom line. Note that the top view refers to the orientation of the image that defines the local coordinate system and not to the facade plane. A numerical evaluation can be found in Tab. 2. The results are based on a weighting using  $\tau_d$  and are averaged from 10 individual repetitions. Our approach performs best in terms of the accuracy of the estimated absolute rotations, both before (upper part) and after bundle adjustment (lower part). The computation times are  $\sim 1.1s$  for the *Herz-Jesu-P25* ( $\sim 235$  relative rotations) and  $\sim 0.4s$  for the *fountain-P11* dataset ( $\sim 49$  relative rotations) using a non-optimised Matlab implementation on a standard desktop computer.

## 5. CONCLUSIONS

In this paper we presented a new approach for a weighted rotation averaging in order to estimate absolute rotations as part of a global image orientation. Our approach is independent on initial values for the unknowns, very robust against outliers in relative rotations and homologous points and reaches good results on synthetic as well as real data that can further be improved performing a subsequent bundle adjustment. An image sequence is considered a graph and we presented a new approach to detect and eliminate outliers in the relative rotations based on simple graph propagation in order to perform iterative least squares averaging. We

|                 | <i>fountain-P11</i> | <i>Herz-Jesu-P25</i> |
|-----------------|---------------------|----------------------|
|                 | $D(\mathbf{R})$     | $D(\mathbf{R})$      |
| This paper      | <b>0.409</b>        | <b>0.556</b>         |
| (1)             | 0.517               | 0.573                |
| This paper (BA) | <b>0.146</b>        | <b>0.167</b>         |
| (1) (BA)        | 0.195               | 0.188                |
| (2)             | 0.420               | 0.348                |

Table 2: Orientation results of the two datasets with respect to the ground truth in comparison with results from recent publications: (1): (Jiang et al., 2013), (2): (Arie-Nachimson et al., 2012). The upper part shows results before, the lower part after performing bundle adjustment.  $D(\mathbf{R})$  is the mean distance to ground truth in  $[\circ]$ .

estimate absolute rotations in the Lie algebra  $\mathfrak{so}(3)$  that is connected with  $SO(3)$  via the exponential map, thus we derive valid rotations accomplishing every necessary constraint. We evaluated three different indicators for a weighting of the relative rotations. Based on synthetic and real data we showed, that the number of homologous points or the distribution of these points in image space are valuable indicators to perform a weighting leading to a significant improvement of the solution. Not only the accuracy itself is improved but also the repeatability of the results. Our results for absolute rotations outperform the solutions of recently published approaches in terms of accuracy. In future work we will investigate further possibilities to include prior knowledge about the quality of the relative rotations into the adjustment. A more sophisticated way would be the variance propagation from homologous points in order to derive the covariance matrix for each relative rotation as prior knowledge. Finally, our approach provides a reliable, accurate and fast way to estimate absolute rotations, suitable also for near real time applications. Thus, future work includes the integration of parts of our approach into on-line compatible applications.

## References

Agarwal, S., Snavely, N., Simon, I., Seitz, S. and Szeliski, R., 2009. Building rome in a day. In: IEEE International Conference 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: International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), IEEE, pp. 81–88.
- Chatterjee, A. and Govindu, V. M., 2013. Efficient and robust large-scale rotation averaging. In: IEEE International Conference on Computer Vision (ICCV), pp. 521–528.
- Dai, Y., Trunpf, J., Li, H., Barnes, N. and Hartley, R., 2010. Rotation averaging with application to camera-rig calibration. In: Asian Conference on Computer Vision (ACCV), Springer, pp. 335–346.
- Enqvist, O., Kahl, F. and Olsson, C., 2011. Non-sequential structure from motion. In: IEEE International Conference on Computer Vision Workshops (ICCV Workshops), 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.
- Govindu, V. M., 2001. Combining two-view constraints for motion estimation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 2, pp. 218–225.
- Govindu, V. M., 2004. Lie-algebraic averaging for globally consistent motion estimation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 1, pp. 684–691.
- Govindu, V. M., 2006. Robustness in motion averaging. In: Asian Conference on Computer Vision (ACCV), Springer, pp. 457–466.
- Hartley, R., Aftab, K. and Trunpf, J., 2011. L1 rotation averaging using the Weiszfeld algorithm. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 1, pp. 3041–3048.
- Hartley, R., Trunpf, J., Dai, Y. and Li, H., 2013. Rotation averaging. *International Journal of Computer Vision* 103(3), pp. 267–305.
- Jiang, N., Cui, Z. and Tan, P., 2013. A global linear method for camera pose registration. In: IEEE International Conference on Computer Vision (ICCV), pp. 481–488.
- Kanatani, K., 1990. *Group-theoretical methods in image understanding*. Vol. 2, Springer.
- Klein, H. and Förstner, W., 1984. Realization of automatic error detection in the block adjustment program pat-m43 using robust estimators. *International Archives of Photogrammetry and Remote Sensing* 25, pp. 234–245.
- Krærup, T., Juhl, J. and Kubik, K., 1980. Gotterdammerung over least squares adjustment. *International Archives of Photogrammetry* 23, pp. 369–378.
- Martinec, D. and Pajdla, T., 2007. Robust rotation and translation estimation in multiview reconstruction. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 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: IEEE International Conference on Computer Vision (ICCV), pp. 3248–3255.
- Reich, M. and Heipke, C., 2014. A global approach for image orientation using lie algebraic rotation averaging and convex  $l_\infty$  minimization. In: ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-3number 1, pp. 265–272.
- 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.
- Snively, N., Seitz, S. M. and Szeliski, R., 2006. Photo tourism: Exploring photo collections in 3d. In: *ACM Transactions on Graphics*, Vol. 25number 3, pp. 835–846.
- 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: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8.
- Weiszfeld, E., 1937. Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tohoku Math. J* 43, pp. 355–386.
- 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.
- Zach, C., Klopschitz, M. and Pollefeys, M., 2010. Disambiguating visual relations using loop constraints. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1426–1433.