

## A GLOBAL APPROACH FOR IMAGE ORIENTATION USING LIE ALGEBRAIC ROTATION AVERAGING AND CONVEX $L_\infty$ MINIMISATION

M. Reich,\* C. Heipke

Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, Germany  
reich@ipi.uni-hannover.de

**KEY WORDS:** image orientation, rotation averaging, convex optimisation, bundle adjustment

### ABSTRACT:

In this paper we present a new global image orientation approach for a set of multiple overlapping images with given homologous point tuples which is based on a two-step procedure. The approach is independent on initial values, robust with respect to outliers and yields the global minimum solution under relatively mild constraints. The first step of the approach consists of the estimation of global rotation parameters by averaging relative rotation estimates for image pairs (these are determined from the homologous points via the essential matrix in a pre-processing step). For the averaging we make use of algebraic group theory in which rotations, as part of the special orthogonal group  $SO(3)$ , form a Lie group with a Riemannian manifold structure. This allows for a mapping to the local Euclidean tangent space of  $SO(3)$ , the Lie algebra. In this space the redundancy of relative orientations is used to compute an average of the absolute rotation for each image and furthermore to detect and eliminate outliers. In the second step translation parameters and the object coordinates of the homologous points are estimated within a convex  $L_\infty$  optimisation, in which the rotation parameters are kept fixed. As an optional third step the results can be used as initial values for a final bundle adjustment that does not suffer from bad initialisation and quickly converges to a globally optimal solution. We investigate our approach for global image orientation based on synthetic data. The results are compared to a robust least squares bundle adjustment. In this way we show that our approach is independent of initial values and more robust against outliers than a conventional bundle adjustment.

### 1. INTRODUCTION

One basic task in photogrammetry is to derive three-dimensional object information and orientation parameters from a set of two-dimensional overlapping images. Most approaches in image orientation are based on a simultaneous estimation of object information and the orientation parameters of the imagery (e. g. bundle adjustment, simultaneous localization and mapping (SLAM) or structure from motion (SfM)). Though in postprocessing all the imagery is available at once, in particular for applications involving image sequences, it is rather common to proceed in a sequential manner: In this case the estimation starts with a certain minimum number of images (often only two or three) and additional images are added iteratively to determine initial values for an overall bundle adjustment. In this process errors are accumulated and hence the final result may suffer from only locally optimised parameters and vary by changing that initial configuration.

An alternative to the sequential procedure is a global two-fold approach: In the first step absolute rotation parameters of each image are estimated from relative rotation estimates. For a set of  $n$  overlapping images there are at least  $n - 1$  and at most  $\frac{n(n-1)}{2}$  different image pairs. In most cases the actual number lies somewhere in between these bounds providing redundant observations for the relative rotations and therefore allowing for an adequate averaging scheme. The second step consists of the estimation of the unknown translation parameters and the object coordinates of homologous points while the rotation parameters are kept constant. Each of the two steps can be simultaneously performed for the whole set of images.

In this paper, we present a new two-step approach for the orientation of a set of overlapping images. Our approach is robust with

respect to outliers and independent of initial values for the unknowns, only the datum has to be fixed. In the first step, absolute rotations  $\mathbf{R}_j$  are computed using estimates for the relative orientations between two images  $\mathbf{R}_{ij}$  that have been derived from tie points (homologous points) via the essential matrices. To take advantage of the redundancy we use all relative estimates available in an iterative block type averaging algorithm in which only one absolute rotation is computed at a time.

In the second step the remaining unknowns, namely the translation parameters of the images and the object coordinates of the homologous points are estimated within a global and convex optimisation. This optimisation is performed by means of an  $L_\infty$  minimisation in a second order cone program.

The remainder of this paper is organised as follows: In section 2 some related work is presented primarily in the field of rotation averaging and global orientation. Section 3, after giving a short introduction into the mathematical field of groups and manifolds, describes single rotation averaging and our algorithm for multiple rotation averaging. The estimation of translation parameters and object coordinates is discussed in section 4. In section 5 we present some results. We summarise and conclude this paper in section 6.

### 2. RELATED WORK

Rotation matrices obey some specific properties, they have a determinant of one and orthogonal columns with unit length. Hence, taking an element-wise average of two rotation matrices most likely will not lead to a valid rotation matrix. Valid averaging of rotations requires some knowledge in mathematical group theory and calculation on manifolds. Rotations as part of the special orthogonal group  $SO(3)$  form a Lie group with a Riemannian manifold structure. This implies that there exists a local Euclidean tangent space, the Lie algebra  $\mathfrak{so}(3)$ , a direct mapping between

\*Corresponding author.

$SO(3)$  and  $\mathfrak{so}(3)$  and a certain metric to perform a proper averaging.

One of the first works exploring the idea of averaging rotations for image orientation is given by Govindu (2001). This paper describes rotation averaging as a least squares estimation using quaternions for representing rotations with residuals being modelled as rotations about an arbitrary axis. The fact that only unit quaternions (quaternions of length one) represent valid rotations is neglected. In a later work the author extends his approach to include Lie group theory (Govindu, 2004). Both, global rotations and translations, derived from essential matrices, are estimated in a single algorithm using a first-order approximation. The structure of the scene is omitted. Fredriksson and Olsson (2013) show how to check whether the estimation of Govindu (2001) reaches the global minimum by analysing the duality gap derived from the Lagrangian dual function.

A fundamental algorithm for our work is the Weiszfeld algorithm, which finds the  $L_1$  mean of a set of points in  $\mathbb{R}^n$  and was published more than 70 years ago (Weiszfeld, 1937) (an English translation is given by Weiszfeld and Plastria (2009)). Hartley et al. (2011) transfer this algorithm to the field of rotation averaging in the Lie algebra. In case of single rotation averaging, meaning the estimation of one absolute rotation from several relative ones, this algorithm is proven to converge to the global optimum. A detailed proof, a broad collection of related work and a mathematical review about rotation averaging and convexity on manifolds is given in Hartley et al. (2013). For multiple rotation averaging they propose a block type algorithm, that estimates only one absolute rotation at the same time and then iterates in the manner of a distributed consensus until convergence. Although the convexity of this optimisation was not proven the authors state that the results demonstrate good performance.

The work of Martinec and Pajdla (2007) proposes to determine rotation averaging by solving a linear system in  $\mathbb{R}^{3 \times 3}$  which will not result in a valid rotation matrix, meaning that the solution is not part of  $SO(3)$ . This is resolved by taking the closest member of  $SO(3)$  with respect to the Frobenius norm. The unknown translation parameters are estimated by means of an  $L_\infty$  minimisation. Hartley et al. (2013) mention that the rotation solution, in general, is not the optimal one. Finally, Chatterjee and Govindu (2013) present a two-fold algorithm. First, a robust initial guess for the global rotations is derived using  $L_1$  optimisation. The result is then improved in a robust least squares estimation with a Huber loss function<sup>1</sup>.

Convex optimisation techniques for geometric imaging problems have become quite popular in recent years. One of the first works is given by Hartley and Schaffalitzky (2004) who analyse  $L_\infty$  optimisation for geometric reconstruction problems. Several works picked up the idea and extended the scope for more geometric problems (Kahl, 2005), several types of convexity (Olsson and Kahl, 2010) and different minimisation algorithms (Ke and Kanade, 2007). A good introduction and an extensive overview to convex optimisation in general can be found in (Boyd and Vandenberghe, 2004).

### 3. ROTATION AVERAGING

In this section we first give an introductory overview about group theory and manifolds. More information on this topic and its relation to image orientation can be found in (Kanatani, 1990). Afterwards we present the single rotation averaging problem and our algorithm for multiple rotation averaging.

<sup>1</sup>the authors call it iterative reweighted least squares

### 3.1 Group theory and manifolds

In the following we will give a brief introduction and a basic definition for the mathematical terms *group* and *manifold*. In algebra a group is defined by a set of elements and an operation that allows to combine elements of the set. Suppose  $A, B$  and  $C$  are elements of a set  $G$ . In order for this set to form a group  $(G, \circ)$ , with operation  $\circ$ , four group axioms must be valid:

1. Closure under operation; every combination of two elements of the group with the associated operation has to result in an element of the group:  $A \circ B = C$ .
2. Associativity:  $A \circ (B \circ C) = (A \circ B) \circ C$ .
3. Identity element; there exist an element  $E$  in  $G$  that does not change other elements of the group when both are combined:  $E \circ A = A \circ E = A$ .
4. Inverse element; for each element in  $G$  there exist an element  $I$  in  $G$  so that the combination of the two results in the identity element:  $I \circ A = A \circ I = E$ .

Rotations  $\mathbf{R}$  can amongst others be represented as matrices with orthogonal columns of unit length and determinant  $\det(\mathbf{R}) = 1$  (orthonormal matrix). They form a group when considering multiplication as operation and  $E$  and  $I$  being the three dimensional identity matrix and the transposed  $\mathbf{R}^T$ , respectively. The name of the group is related to the main characteristics of rotations: *special orthogonal group*,  $SO(3)$ . *Orthogonal* is related to the matrix being orthogonal, *special* stands for the determinant equal to one,  $3$  specifies the matrix dimension.

A manifold is a topological space, hence a set with a topological structure that allows the formulation of a metric. Manifolds obey Euclidean geometry only locally, but not necessarily also globally. A manifold is a *Riemannian* manifold if it is differentiable (allows the computation of a tangent space) and induces a metric tensor<sup>2</sup> (allows the definition of lengths and angles in the tangent space). Basically, the latter follows from the former. Each point of a *Riemannian* manifold has a certain neighbourhood in which the Euclidean geometry is approximately valid.

An intuitive example is the surface of a 2-sphere (one can think of the earth's surface). Globally the surface is not Euclidean, for instance the sum of the three angles of a spherical triangle exceeds  $180^\circ$  and the shortest path between two points is not a straight line but a geodesic. The surface can be mapped to Euclidean  $\mathbb{R}^2$  but the larger the mapped region, the larger the distortion.

Both mathematical concepts, group and Riemannian manifold, are connected in a *Lie group*. A group is a Lie group if its operation and the computation of the inverse element are compatible with the structure of the manifold, hence both lead to continuous and differentiable functions. Consequently, we can state that  $SO(3)$  is a Lie group, since the multiplication of rotation matrices is continuous and differentiable. Every Lie group is connected to a Lie algebra, the tangent space at the identity element (we will write  $\mathfrak{so}(3)$  for the Lie algebra of  $SO(3)$ ). The Lie algebra obeys the Euclidean geometry and hence allows for an adequate averaging. The matrix logarithm<sup>3</sup>  $\log : SO(3) \rightarrow \mathfrak{so}(3)$  connects the Lie group and the Lie Algebra. An inverse mapping is given by the exponential map  $\exp : \mathfrak{so}(3) \rightarrow SO(3)$ .

<sup>2</sup>A generalisation of the dot product in Euclidean space

<sup>3</sup>The matrix logarithm is computed by taking the logarithm of the eigenvalues of the matrix and re-assembling.

### 3.2 Single rotation averaging

The single rotation averaging problem comprises computing a valid rotation given several estimates which are assumed to be affected by noise. We have seen that rotations form an  $SO(3)$ , a Lie group with a Riemannian manifold structure. Normally, when we speak about averaging, this is meant to be done in Euclidean geometry. Hence, in the following we make use of the fact, that rotations can be mapped to their local Euclidean tangent space and vice versa. Hartley et al. (2011) and Hartley et al. (2013) name three different metrics for defining distances between elements of  $SO(3)$ : The geodesic, the chordal and the quaternion metric which basically differ in scale. In our approach we make use of the geodesic metric, which is also called angular metric since it defines the angle between two rotations  $\mathbf{R}_1$  and  $\mathbf{R}_2$  and appears to be the most intuitive one:

$$d(\mathbf{R}_1, \mathbf{R}_2) = \|\log(\mathbf{R}_1^T \mathbf{R}_2)\|_2. \quad (1)$$

Note that the matrix logarithm in (1) maps the combined rotation from  $SO(3)$  to  $\mathfrak{so}(3)$  and that the norm is the standard Euclidean norm in  $\mathbb{R}^3$ . With the distance measure defined in (1) we are interested in finding a rotation  $\mathbf{R}$  with a minimum distance to every estimate  $\mathbf{R}_i$  with respect to some norm. The  $L_p$  cost function for a set of  $n$  rotations reads as

$$D(\mathbf{R}) = \sum_i^n d(\mathbf{R}_i, \mathbf{R})^p. \quad (2)$$

In our work we analyse both, the  $L_1$  and the  $L_2$  optimisation. For the  $L_1$  optimisation we use the adapted Weiszfeld algorithm presented in Hartley et al. (2011). Both the  $L_1$  and the  $L_2$  algorithms are gradient descent algorithms. The  $L_2$  optimisation can start with an arbitrary rotation, for instance one of the initial estimates to be averaged, whereas the gradient of the  $L_1$  cost function is not defined at any of the estimates. Therefore, we use the  $L_2$  solution after only one iteration providing a rotation different from any of the estimates as initial value for the Weiszfeld algorithm.

The  $L_1$  minimisation, also known as the geometric median, has a breakdown point of 0.5 and offers a higher robustness than the  $L_2$  optimisation. In  $L_2$  high residuals lead to a much higher cost and hence to a stronger influence of outliers. On the other hand  $L_2$  optimisation leads to the best unbiased estimation in Euclidean space. In order to achieve a high robustness in an  $L_2$  optimisation we use a reweighted averaging algorithm based on the weighting function

$$w(i) = \begin{cases} \frac{2}{z(i)^2} & \text{if } z(i) \leq k \\ \frac{1}{k \cdot z(i)} - \frac{2}{k^2} & \text{else,} \end{cases} \quad (3)$$

with  $z(i) = \frac{d(\mathbf{R}_i, \hat{\mathbf{R}})}{D(\hat{\mathbf{R}})}$

$i = 1, \dots, n,$

which is the inverse Huber loss function (Huber et al., 1964) with tuning constant  $k^4$  and  $\hat{\mathbf{R}}$  being the current rotation average. Thus, while good estimates have a quadratic influence on the results, bad estimates only have a linear one.

<sup>4</sup> $k$  is selected by means of a proper quantile of the normal distribution.

### 3.3 Multiple rotation averaging

In the last section we have seen how one rotation can be averaged from several estimates using different cost functions. In this section we want to extend this problem to a network of  $m$  estimated relative rotations that is to be solved for  $n \leq m$  absolute rotations. According to (Chatterjee and Govindu, 2013) one can imagine this network as a graph with nodes being absolute rotations connected by edges depicting relative estimates. Rotation matrices are composed by multiplication and they can be chained. Hence our functional model reads as:

$$\mathbf{R}_j = \mathbf{R}_{i,j} \mathbf{R}_i, \quad \forall (i, j) \in N = 1, \dots, n; i \neq j. \quad (4)$$

In general this equation does not hold, since every estimate most likely is affected by noise. Hence, we formulate the  $L_p$  cost function for the multiple rotation averaging problem in the manner of equation (2) and according to Hartley et al. (2011), Chatterjee and Govindu (2013):

$$D(\mathbf{R}_1, \dots, \mathbf{R}_n) = \sum_{(i,j) \in N} d(\mathbf{R}_{i,j} \mathbf{R}_i, \mathbf{R}_j)^p. \quad (5)$$

Again,  $d(\mathbf{R}_{i,j} \mathbf{R}_i, \mathbf{R}_j)^p$  stands for the geodesic metric defined in (1). Equation (4) implies the need of one fixed absolute rotation. Without loss of generality we set the rotation of the first image to the identity matrix to fix the datum.

In the following we will describe our procedure, which is similar in style to the one outlined in Hartley et al. (2011) but in addition contains a scheme for outlier detection. Basically, the algorithm can be divided into two stages:

1. initialisation and outlier detection
2. iterative averaging.

The first stage deals with all  $n$  absolute rotations, for which an average has to be computed, separately. This step serves the initialisation of the graph, thus one can think of assigning a rotation to each node of the graph. The principle is outlined in figure 1. Starting from the first image  $i_1$  we select each edge corresponding to a relative rotation from or to another absolute rotation and propagate the rotation of all the involved nodes  $i_2, i_3$  and  $i_4$  according to (4). In the beginning this is quite intuitive since only for the first node a rotation has been assigned. Propagation becomes more challenging when a node already exhibits an assigned rotation (figure 1b). Before these two estimates are averaged (e. g.  $\mathbf{R}_3$  and  $\mathbf{R}_{23} \mathbf{R}_2$ ), we check them for similarity using the distance measure defined in (1). Based on this measure we decide whether or not one of the estimates is an outlier. If so, of course, we cannot yet decide which of the two is the bad one. Therefore, we defer our decision, keeping both estimates for the specific node and proceed with the following relative rotation estimate, if available. Since relative rotations are estimated independently, we argue that, if another relative rotation from this particular node ( $i_2$ ) leads to a high discrepancy in the estimates (e. g.  $\mathbf{R}_4$  and  $\mathbf{R}_{24} \mathbf{R}_2$ ), the rotation of this node,  $\mathbf{R}_2$ , can be considered an outlier. In this case the current solution for this node is discarded and, redundancy provided, computed with another relative estimate at a later stage. This also means, that propagation

from this node is no longer possible since it exhibits no valid rotation. Propagation from this node is only repeated after it has been assigned another rotation. All the involved nodes (in this case  $i_3$  and  $i_4$ ) keep their assigned rotation without taking the relative estimates from the discarded node ( $i_2$ ) into account. If there is no further evidence for  $\mathbf{R}_2$  being an outlier (e. g.  $d(\mathbf{R}_4, \mathbf{R}_{24}\mathbf{R}_2)$  is small) we keep the estimates for the rotation that could not yet be averaged and decide when a third estimate comes from another node.

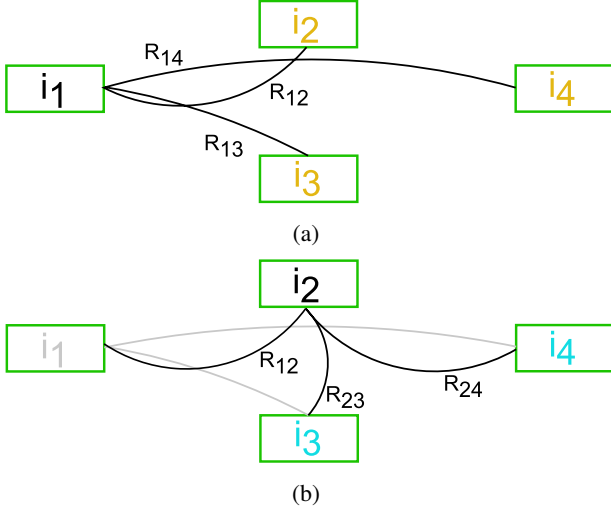


Figure 1: Graph structure during multiple rotation averaging depicted for two phases of the first stage. (1a) shows propagation from image  $i_1$  (black) in which all other nodes receive a rotation (orange). In (1b) propagation starts from image  $i_2$ .  $i_1$  does not receive a rotation since it is fixed (grey).  $i_3$  and  $i_4$  exhibit an assigned value and underlie outlier detection and averaging (turquoise).

If we do not qualify one estimate as an outlier the two estimates are averaged using the cost function described in (2). This procedure is repeated for every node of the graph. With each repetition our outlier detection becomes more stable because more and more estimates are involved in propagation supporting the rotation of each node.

In the first stage for all  $n$  rotations to be averaged initial values are determined. The second stage is carried out for a certain number of iterations steadily improving the overall solution. For each node every connected neighbour is consulted to propagate a rotation. Again, these estimates are averaged using (2). Within the iteration the overall solution is found in a manner similar to relaxation labelling (Kittler and Illingworth, 1985). This algorithm is not convex and may lead to local minima (Hartley et al., 2011). Nevertheless, the authors state that it gives excellent results even on large datasets.

#### 4. TRANSLATION AND STRUCTURE ESTIMATION

In section 3 we examined how to compute global rotations using the redundancy provided by  $m \geq n$  relative rotations. In this section we show the convex estimation of translation parameters and the 3D coordinates of the tie points. The convex formulation of the multiview reconstruction problem with known rotations has been studied in many publications in recent years (Hartley and Schaffalitzky, 2004), (Ke and Kanade, 2007), (Olsson and Kahl, 2010), (Kahl, 2005). All these works examine the quasiconvexity

of an  $L_\infty$  cost function for several multivision problems like spatial intersection (sometimes also called triangulation), spatial resection and homography estimation in addition to our problem. A quasiconvex function has only one global minimum but, contrary to a convex function, may have negative curvature and more than one stationary point (Boyd and Vandenberghe, 2004). Hence a gradient  $\nabla f(x) = 0$  does not necessarily stand for a global minimum. Therefore, a quasiconvex problem cannot be solved by a gradient descent method. However, there exist methods to solve quasiconvex problems, for instance bisection algorithms (Ke and Kanade, 2007) that subdivide the quasiconvex problem into several convex problems.

Hartley and Schaffalitzky (2004) show that the reprojection error in image space is a quasiconvex function with respect to the point movement in object space. Generally in image orientation one is interested in simultaneously minimising all reprojection errors of the involved tie points, but quasiconvexity is not preserved under summation. Instead using the  $L_\infty$  norm  $\|f(x)\|_\infty = \max f(x)$ , thus minimising the maximum residual, preserves quasiconvexity allowing for a convex estimation of the unknowns. It should be noted that since  $L_\infty$  optimisation is sensitive to outliers using this norm assumes outlier-free data.

We can write the Euclidean reprojection error as:

$$r_{i,j}(\mathbf{X}_j, \mathbf{X}_{0,i}) = \left\| \begin{pmatrix} -c \frac{\mathbf{r}_{i,1}^T (\mathbf{X}_j - \mathbf{X}_{0,i})}{\mathbf{r}_{i,3}^T (\mathbf{X}_j - \mathbf{X}_{0,i})} + x_0 - x_{i,j} \\ -c \frac{\mathbf{r}_{i,2}^T (\mathbf{X}_j - \mathbf{X}_{0,i})}{\mathbf{r}_{i,3}^T (\mathbf{X}_j - \mathbf{X}_{0,i})} + y_0 - y_{i,j} \end{pmatrix} \right\|_2, \quad (6)$$

$i = 1 \dots n, j = 1 \dots p,$

with  $\mathbf{X}_j$  and  $\mathbf{X}_{0,i}$  being the unknown object coordinate of the  $j$ th tie point and the projection centre of the  $i$ th image.  $[x_{i,j}; y_{i,j}]$  are the observed image coordinates of point  $j$  in image  $i$ . The interior orientation elements  $(c, x_0, y_0)$  and the columns of the rotation matrix of image  $i$ ,  $\mathbf{R}_i = [\mathbf{r}_{i,1}, \mathbf{r}_{i,2}, \mathbf{r}_{i,3}]$  are constant. We can extend (6) to

$$r_{i,j}(\mathbf{X}_j, \mathbf{X}_{0,i}) = \left\| \begin{pmatrix} \frac{(-c \mathbf{r}_{i,1}^T + (x_0 - x_{i,j}) \mathbf{r}_{i,3}^T) (\mathbf{X}_j - \mathbf{X}_{0,i})}{\mathbf{r}_{i,3}^T (\mathbf{X}_j - \mathbf{X}_{0,i})} \\ \frac{(-c \mathbf{r}_{i,2}^T + (y_0 - y_{i,j}) \mathbf{r}_{i,3}^T) (\mathbf{X}_j - \mathbf{X}_{0,i})}{\mathbf{r}_{i,3}^T (\mathbf{X}_j - \mathbf{X}_{0,i})} \end{pmatrix} \right\|_2, \quad (7)$$

and further to

$$r_{i,j}(\mathbf{X}_j, \mathbf{X}_{0,i}) = \frac{\| \mathbf{A}_{i,j} \begin{pmatrix} \mathbf{X}_j \\ \mathbf{X}_{0,i} \end{pmatrix} \|_2}{\mathbf{c}_i^T \begin{pmatrix} \mathbf{X}_j \\ \mathbf{X}_{0,i} \end{pmatrix}}, \quad (8)$$

stacking the two unknown vectors and with

$$\mathbf{A}_{i,j} = \begin{pmatrix} -\mathbf{a}_{i,j} & \mathbf{a}_{i,j} \end{pmatrix}_{2 \times 6},$$

$$\mathbf{a}_{i,j} = \begin{pmatrix} c \cdot \mathbf{r}_{i,1}^T + (x_0 - x_{i,j}) \mathbf{r}_{i,3}^T \\ c \cdot \mathbf{r}_{i,2}^T + (y_0 - y_{i,j}) \mathbf{r}_{i,3}^T \end{pmatrix}_{2 \times 3},$$

$$\mathbf{c}_i^T = \begin{pmatrix} \mathbf{r}_{i,3}^T & -\mathbf{r}_{i,3}^T \end{pmatrix}_{1 \times 6}.$$

(8) is a quasiconvex function on the convex set  $\mathbf{c}^T \begin{pmatrix} \mathbf{X}_j \\ \mathbf{X}_{0,i} \end{pmatrix} \geq 0$  requiring the points to lie in front of the cameras (cheirality constraint). The resulting  $L_\infty$  cost function then reads as:

$$D(\mathbf{X}, \mathbf{X}_0) = \max_{i,j} r_{i,j}(\mathbf{X}, \mathbf{X}_0). \quad (9)$$

In minimising (9) we follow Ke and Kanade (2007) and formulate the following convex minimisation problem:

$$\begin{aligned} & \text{minimise} && \epsilon \\ & \text{subject to} && \left\| \mathbf{A}_{i,j} \begin{pmatrix} \mathbf{X}_j \\ \mathbf{X}_{0,i} \end{pmatrix} \right\|_2 - \mu \left( \mathbf{c}_i^T \begin{pmatrix} \mathbf{X}_j \\ \mathbf{X}_{0,i} \end{pmatrix} \right) \leq \epsilon, \\ & && \mathbf{c}_i^T \begin{pmatrix} \mathbf{X}_j \\ \mathbf{X}_{0,i} \end{pmatrix} \geq 0, \\ & && i = 1 \dots n, j = 1 \dots p. \end{aligned} \quad (10)$$

The parameter  $\mu$  is the reprojection error  $r_{i,j}(\mathbf{X}_j, \mathbf{X}_{0,i})$  that is kept fixed and set to a reasonable value. For each image  $i$  and each point  $j$  the first constraint defines a second order cone, which is why (10) is called a second order cone program (Boyd and Vandenberghe, 2004).  $\mu$  can geometrically be thought of as defining the diameter in image space of each convex cone of feasible solutions. The objective function  $\epsilon$  is an affine function with respect to the reprojection error and  $\epsilon \leq 0$  holds if a solution inside the feasible set defined by all cones is found. The datum has to be fixed, which means that three translational parameters and the scale have to be defined. An intuitive way is to locate the origin of the coordinate system in the projection centre of the first image and to fix the  $X$ -coordinate of the second projection centre. Hence, (10) extends to

$$\begin{aligned} & \text{minimise} && \epsilon \\ & \text{subject to} && \left\| \mathbf{A}_{i,j} \begin{pmatrix} \mathbf{X}_j \\ \mathbf{X}_{0,i} \end{pmatrix} \right\|_2 - \mu \left( \mathbf{c}_i^T \begin{pmatrix} \mathbf{X}_j \\ \mathbf{X}_{0,i} \end{pmatrix} \right) \leq \epsilon, \\ & && \mathbf{c}_i^T \begin{pmatrix} \mathbf{X}_j \\ \mathbf{X}_{0,i} \end{pmatrix} \geq 0, \\ & && \mathbf{X}_{0,1} = \mathbf{0}_{3 \times 1}, \\ & && \mathbf{X}_{0,2,(1)} = 1, \\ & && i = 1 \dots n, j = 1 \dots p. \end{aligned} \quad (11)$$

Note that in (11)  $\mathbf{X}_{0,2,(1)}$  is set equal to 1 but in general can be set to an arbitrary value.

## 5. EXPERIMENTAL RESULTS

In this section we present some of our experimental results based on synthetic data. First we analyse our robust  $L_2$  single rotation averaging algorithm. Whereas in an outlier free dataset  $L_2$  optimisation generally performs best, this is not the case in the presence of gross errors. Out of one rotation matrix we generated five estimates by adding random noise to the Euler angles. In a Monte-Carlo simulation we computed 1000 solutions averaging the five rotation estimates using (2) for  $L_1$  and robust  $L_2$  optimisation based on (3). We successively increased the noise level and

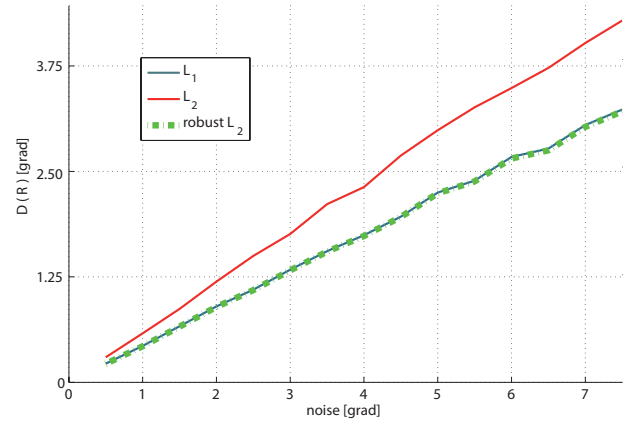


Figure 2: Distance according to (1) between the true rotation and an averaged solution ( $L_1$ ,  $L_2$  and weighted  $L_2$ ) out of five independent rotation estimates with 10% outlier rate.

assumed every tenth rotation estimate to be an outlier with noise being 5 times higher than the current noise level. The results are shown in figure 2. As one can see the  $L_2$  solution is affected by the outliers more significantly and thus produces worse results than both,  $L_1$  and robust  $L_2$  optimisation which perform equally well.

Because the  $L_1$  optimisation has to be initialised and is computationally is more expensive in the following experiments we make use of our robust  $L_2$  algorithm.

The second experiment serves as an illustration of the quality of our combined approach using multiple rotation averaging and convex estimation of translations and structure<sup>5</sup>. We created a setting with a circular arrangement of images. Twelve images located in the  $XZ$ -plane are rotated through  $n \cdot \frac{\pi}{6}$  about the  $Y$ -axis ( $n = 0 \dots 11$ ); 100 object points are uniformly distributed within a cuboid in the centre. Each point is observed in all images and we assume error-free matches. We argue that gross errors within matching most likely are detected during the subsequent estimation of the relative orientation parameters via essential matrices for instance by means of a robust RANSAC algorithm. All the observations are affected by random noise, normally distributed with  $\sigma = 5$  pixels. The graph that depicts the distribution of relative estimates for multiple rotation averaging (see section 3.3) is fully connected.

Figure 3 shows the configuration of images (green image planes) and object points (green squares). The coloured dots and the red image planes depict the results using different algorithms. We derived initial values by computing the essential matrix of the first image pair (the topmost image and the one to its right at (4, 0, 2)), spatial intersection of the tie points and the computation and decomposition of projection matrices for the rest of the images. In figure 3a, that shows the initial configuration, one can see, that the first image pair is approximated quite accurately whereas the images at the opposite side of the circle differ significantly from the ground truth. The reason is that the projection matrices are computed using the initial object coordinates intersected from the first image pair. Due to the small base length the points exhibit a comparatively large uncertainty in depth. The results of our algorithm can be seen in figure 3c. Note that no initial values are needed. In comparison with figure 3b which shows results achieved without multiple rotation averaging using only the first initial guess

<sup>5</sup>We used CVX, a Matlab package for specifying and solving convex programs (Grant and Boyd, 2014), (Grant and Boyd, 2008)

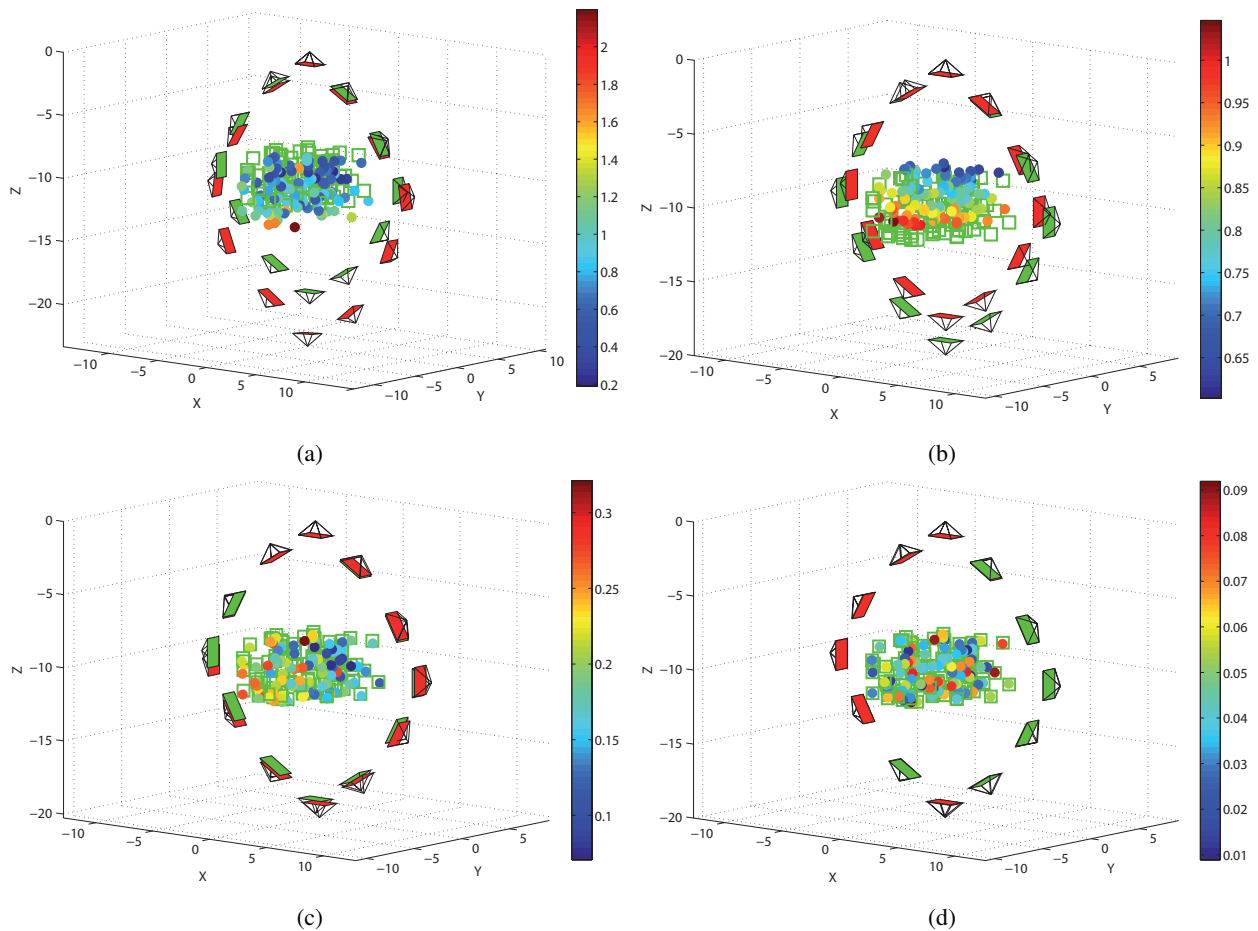


Figure 3: Circular arrangement of images and randomly distributed object points. The green images and the green squares depict the ground truth. The red images and the coloured points show the results of the following stages: initial values from projective geometry (3a), solution of convex estimation of translations and object points without rotation averaging (3b), solution after multiple rotation averaging and convex estimation of translations and object points (3c) and solution of a robust bundle adjustment (3d) using the initial values depicted in (3a). The points are coloured with respect to their distance to the ground truth. Note that the color scale is different in every stage.

for absolute rotations and hence neglecting the redundancy of the relative estimates, one can see a large improvement. Because the rotations are kept fixed during the estimation of the translations and object coordinates, there exist one globally optimal solution for each set of rotations. In figure 3d one can see the results of the bundle adjustment. Clearly, since an  $L_2$  minimisation is performed, this method achieves the best results with respect to the Euclidean distance to the ground truth.

Numerical results for the RMS-values of the unknowns are given in table 1. Performing rotation averaging leads to a significant improvement of nearly 300% of the RMS of the projection centres and the object coordinates with respect to the ground truth. The improvement of the rotational parameters amounts to a factor of 2. Still, the accuracy of our algorithm is far away from the accuracy of the robust bundle adjustment. However, the result of our approach can serve as a good initialisation for a final bundle adjustment.

In another experiment we demonstrate the robustness of our method. We generated initial values for the orientation parameters and object coordinates by adding several combinations of random noise to the ground truth. In a Monte-Carlo simulation we ran the robust bundle adjustment 20 times for each combination of initial value accuracy. We then took the median from the 20

RMS-values of the object coordinates, the translation and the rotation parameters, respectively. Exemplary results for the RMS of the translation parameters are depicted in figure 4. We draw samples with an accuracy from 0.025 to 0.125 base lengths [BL] for initial translations and object coordinates and from 0.5 to 5 [grad]<sup>6</sup> for initial rotations. One can see clearly that the worse the accuracy of the initial values the worse the result of the bundle adjustment. Note that we did not change the noise of the observations nor the observations during the simulation. This fact also manifests itself in the constant standard deviation of our approach (green and orange planes). Thus, due to poor initialisation the bundle adjustment may converge into only locally optimal solutions.

## 6. CONCLUSION

In this paper we have described a new robust two-step approach for image orientation that is independent of initial values for the unknown parameters and most likely converges to a global optimum with respect to the  $L_\infty$  norm. In the first step we make use of the redundant overlap of the images and compute absolute rotations from relative estimates. Our robust algorithm for multiple rotation averaging is able to detect outliers in terms of gross

<sup>6</sup>noise on the Euler angles



	<b>initial values</b>	<b>convex opt. no rotation av.</b>	<b>convex opt. rotation av.</b>	<b>robust bundle adj.</b>
$RMS_{\mathbf{X}_0}$	1.4816	0.9726	0.2512	0.0765
$RMS_{\mathbf{R}}[grad]$	7.4064	2.3305	1.2728	0.4484
$RMS_{\mathbf{X}}$	0.8376	0.8259	0.1867	0.0456

Table 1: RMS-values of the translation parameters  $\mathbf{X}_0$ , the rotations  $\mathbf{R}$  (using (1)) and the object coordinates  $\mathbf{X}$  for the three methods with respect to the ground truth.

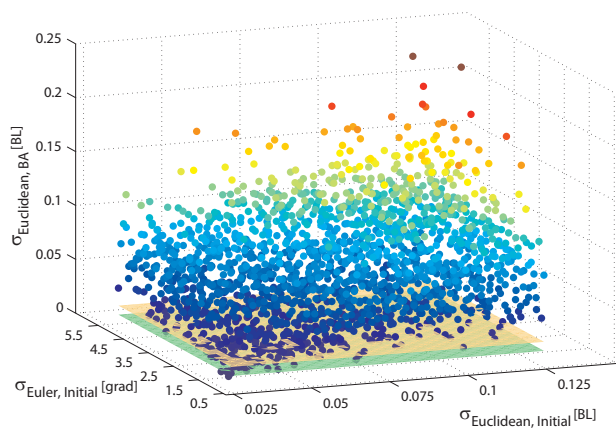


Figure 4: Median of the RMS of the translation parameters after a Monte-Carlo simulation of an incremental robust bundle adjustment using different combinations of initial value accuracies (dots, coloured with respect to the distance to the ground truth). The planes depict the standard deviation of the translation parameters of our approach (orange) and our approach plus a final bundle adjustment (green).

errors in the relative rotation estimates, given that redundant information exist. Translation parameters and object coordinates are estimated using a convex  $L_\infty$  minimisation providing a globally optimal solution for a certain set of constant rotations. A subsequent and quickly converging bundle adjustment, that does not suffer from poor initialisation, is then able to compute a globally optimal  $L_2$  solution. In our rather simple experiments we could show that poor initialisation may lead to only locally optimal results in bundle adjustment. This stresses the need for a highly robust method such as the one we presented in this paper.

There are some aspects that are not dealt with in this work and need further examination. First, we only experimented with synthetic data, thus we need to evaluate our approach based on real data and more complex image configurations. Second, we could not yet define a proper stopping criterion for our multiple rotation averaging algorithm, thus we used a fixed number of iterations. This comes along with a rather slow convergence rate, so in future work we will investigate improvements. Furthermore, we want to find a better way to detect relative rotation outliers, for instance based on the overall standard deviation of the relative estimates which depends on image overlap and the distribution of homologous points in the images. In order to provide good relative estimates, a more robust way for generating the relative rotations by means of the trifocal tensor instead of essential matrices can be analysed. We also want to further evaluate the level of redundancy needed for rotation averaging to provide results not affected by outliers for the subsequent convex  $L_\infty$  optimisation.

What we did not mention so far, because it is not the focus of this paper, is the high computational effort of the  $L_\infty$  minimisation. This can be a problem especially for large image blocks. A

solution might be to reformulate the minimisation problem as a linear program as proposed by Ke and Kanade (2007) which can be solved faster. A difference in the quality of the results also has to be analysed in future work.

## References

- Boyd, S. P. and Vandenberghe, L., 2004. Convex Optimization. Cambridge University Press.
- Chatterjee, A. and Govindu, V. M., 2013. Efficient and robust large-scale rotation averaging. In: IEEE International Conference on Computer Vision, 2013. ICCV 2013., pp. 521–528.
- Fredriksson, J. and Olsson, C., 2013. Simultaneous multiple rotation averaging using lagrangian duality. In: ACCV 2012, Springer, pp. 245–258.
- Govindu, V. M., 2001. Combining two-view constraints for motion estimation. In: IEEE Conference on Computer Vision and Pattern Recognition, 2001. CVPR 2001., Vol. 2, IEEE, pp. II–218.
- Govindu, V. M., 2004. Lie-algebraic averaging for globally consistent motion estimation. In: IEEE Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004., Vol. 1, IEEE, pp. I–684.
- Grant, M. and Boyd, S., 2008. Graph implementations for nonsmooth convex programs. In: V. Blondel, S. Boyd and H. Kimura (eds), Recent Advances in Learning and Control, Lecture Notes in Control and Information Sciences, Springer-Verlag Limited, pp. 95–110. [http://stanford.edu/~boyd/graph\\_dcp.html](http://stanford.edu/~boyd/graph_dcp.html).
- Grant, M. and Boyd, S., 2014. CVX: Matlab software for disciplined convex programming, version 2.1. <http://cvxr.com/cvx>.
- 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 2011, IEEE, pp. 3041–3048.
- Hartley, R. and Schaffalitzky, F., 2004.  $L_\infty$  minimization in geometric reconstruction problems. In: IEEE Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004., Vol. 1, pp. I–504–I–509 Vol.1.
- Hartley, R., Trunpf, J., Dai, Y. and Li, H., 2013. Rotation averaging. International Journal of Computer Vision 103(3), pp. 267–305.
- Huber, P. J. et al., 1964. Robust estimation of a location parameter. The Annals of Mathematical Statistics 35(1), pp. 73–101.
- Kahl, F., 2005. Multiple view geometry and the  $L_\infty$ -norm. In: IEEE International Conference on Computer Vision, 2005. ICCV 2005., Vol. 2, IEEE, pp. 1002–1009.
- Kanatani, K., 1990. Group-theoretical methods in image understanding. Vol. 2, Springer-Verlag Berlin.

- Ke, Q. and Kanade, T., 2007. Quasiconvex optimization for robust geometric reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29(10), pp. 1834–1847.
- Kittler, J. and Illingworth, J., 1985. Relaxation labelling algorithms - a review. *Image and Vision Computing* 3(4), pp. 206–216.
- Martinec, D. and Pajdla, T., 2007. Robust rotation and translation estimation in multiview reconstruction. In: *IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR 2007.*, IEEE, pp. 1–8.
- Olsson, C. and Kahl, F., 2010. Generalized convexity in multiple view geometry. *Journal of Mathematical Imaging and Vision* 38(1), pp. 35–51.
- Weiszfeld, E., 1937. Sur le point pour lequel la somme des distances de  $n$  points donnés est minimum. *Tohoku Math. J* 43(355-386), pp. 2.
- 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.