

# Solving Bundle Block Adjustment by Generalized Anisotropic Procrustes Analysis

Andrea Fusiello\*, Fabio Crosilla

Dipartimento di Ingegneria Elettrica, Gestionale e Meccanica,  
University of Udine, Via Delle Scienze, 208 - 33100 Udine, Italy

Dipartimento di Ingegneria Civile e Architettura,  
University of Udine, Via Delle Scienze, 208 - 33100 Udine, Italy

---

## Abstract

The paper presents a new analytical tool to solve the classical photogrammetric bundle block adjustment. The analytical model is based on the generalized extension of the anisotropic row-scaling Procrustes analysis, that has been recently proposed by the same authors to solve the image exterior orientation problem. The main advantage of the method is given by the fact that the problem solution does not require any approximate value of the unknown parameters, nor any linearization procedure. Moreover, the algorithm is exceedingly simple to describe and easy to implement. Empirical results indicate that a zero-information initialization of the iterative relaxation procedure leads almost always to the correct final least squares solution. Experiments confirm the accuracy of the proposed method, when compared to the results obtained by applying a classical photogrammetric bundle block adjustment.

**Keywords:** photogrammetric bundle block adjustment, anisotropic row-scaling Procrustes analysis, alternating least squares, block relaxation

---

## 1. Introduction

Bundle block adjustment is the most classical analytical problem in photogrammetry. Great effort was given to the problem solution since the middle of the last century (Baetsle, 1956; Brown, 1958; Ackermann, 1962; Cunietti, 1968, e.g.). An exhaustive synthesis of the analytical developments carried out in this fundamental field of photogrammetry can be found, for instance, in Triggs et al. (2000). The standard formulation requires the linearization of the collinearity equations and the satisfaction of the least squares principle for the equation residuals. Some additional unknown terms can be considered for each equation in order to calibrate the camera for some systematic error terms or for simultaneously estimating the image interior orientation parameters. Robust least squares solutions have been also proposed in the literature in order to reduce the influence of outliers (Zhang et al., 2006, e.g.).

According to Triggs et al. (2000), the most significant bundle block adjustment paradigmatic enhancements in chronological sequence, are:

1. recursive partitioning by Gyer (1967) and Brown (1968) that led to the modern sparse matrix techniques;
2. S transformations and criterion matrices by Baarda (1973), that allowed the correct estimate of the network degrees of freedom and the uncertainty modeling in the adjustment process;

3. photogrammetric precision and reliability over-parametrization and model choice by Gruen (1980) and Foerstner (1985), that opened the way to modern robust statistics and model selection in photogrammetry;
4. “geometrically constrained multiphoto and globally enforced least squares matching” by Gruen and Baltzavias (1986), that introduced the so called image-based matching technique procedures.

In spite of these fundamental steps in the methodological development of bundle adjustment, the common underlying scheme, based on the least-squares solution of a large non-linear system of equations, has been the same since its origins. In this paper we propose instead a new analytical bundle block solution method rooted in the framework of orthogonal Procrustes analysis, in particular focusing on its *generalized anisotropic* variant. The main advantage of the method is that – upon convergence – it furnishes a least squares solution without any linearization of the original equations, and without any approximate value of the unknown parameters and of the tie-points 3-D coordinates.

Recently, the authors of this paper applied the anisotropic row-scaling Procrustes analysis technique to successfully solve the exterior orientation problem of one image (Garro et al., 2012). The process is carried out by an iterative relaxation of the unknown translation, rotation and anisotropic scaling of each image point.

This paper, in particular, extends the same relaxation

---

\*Corresponding author

Email addresses: [andrea.fusiello@uniud.it](mailto:andrea.fusiello@uniud.it) (Andrea Fusiello), [fabio.crosilla@uniud.it](mailto:fabio.crosilla@uniud.it) (Fabio Crosilla)

procedure by considering also the presence of unknown tie-points imaged to some or all exposures. Their role is to constraint the different images one to each other in order to iteratively update their approximate exterior orientation parameters until a global convergence of the entire block.

The method introduced in this paper represents a further extension of the generalized (isotropic) Procrustes analysis, already applied for the least squares registration of photogrammetric 3-D models (Crosilla and Beinat, 2002, e.g.), where just one isotropic scale factor is required for all the points of the same model.

The new method falls in the wider structure-from-motion family. Its closest neighbours are the *iterative factorization* methods and the *global motion-first* methods.

Iterative factorization methods (Sturm and Triggs, 1996; Heyden, 1997; Oliensis, 1999; Oliensis and Hartley, 2007) yield a projective model from multiple images by a two step iteration (a block relaxation, in fact), where in one step a measurement matrix, containing image points coordinates, is factorized with SVD, and in the subsequent step the depths of the points are computed, assuming all the other parameters fixed. This bears some resemblance of the scheme described in this paper, which however, deals with calibrated cameras (i.e., known interior orientation) and outputs a Euclidean model<sup>1</sup> instead of a projective one. Moreover, unlike these algorithms, our method does not require all points to be visible in all views.

The issue of visibility in matrix factorization methods can be side-stepped by matrix completion techniques, exploiting the low rank of the measurement matrices (Brand, 2002; Kennedy et al., 2013; Hartley and Schaffalitzky, 2003), or by providing additional information. Indeed, Kaucic et al. (2001); Rother and Carlsson (2002), describe algorithms based on SVD for the projective modeling from multiple perspective views, based on having four points on a reference plane visible in all views. Unlike iterative factorization ones, these algorithms does not require all points to be visible in all views and are also direct. If three orthogonal vanishing points are specified in addition, the model can be upgraded to Euclidean. Hartley and Schaffalitzky (2003) scheme, in particular, can be contrasted with ours, since its iteration can be interpreted as linearly solving alternately for camera matrices and 3D points until their product converges to the measurement matrix.

Projective methods, though, respond to a practical situation (i.e., unknown interior orientation), which is different from the one addressed in this paper. Moreover, we do not put any constraint on the input (like having four points on a reference plane visible in all views).

Global motion-first methods share a common scheme: they start from known interior orientation, compute epipolar or trifocal geometry which results in relative rotations and relative translations (up to a scale). Then solve the *rotation registration* or *rotation averaging* (Hartley et al.,

2013) problem that gives the rotational component of the cameras orientation; this problem, if one ignores outliers, can be solved directly by eigen-decomposition of a matrix (Martinec and Pajdla, 2007; Arie-Nachimson et al., 2012). Finally, camera location is solved (a.k.a. *translation registration*) by a variety of direct/iterative methods, including solving a linear system of equations (Kraus, 1997, Sec. 4.1), (Arie-Nachimson et al., 2012; Jiang et al., 2013), eigen-decomposition (Brand et al., 2004), linear programming (Moulou et al., 2013), Second Order Cone programming (Kahl and Hartley, 2008; Martinec and Pajdla, 2007), non-linear least squares (Wilson and Snavely, 2014).

Since some of these methods are direct and the scheme proposed in this paper is iterative, they can be considered superior from this point of view. However, they minimize algebraic residuals, which in certain cases are based only on the orientations (Brand et al., 2004, e.g.), ignoring the 3D points until the final intersection. On the other hand, our method minimizes a *geometric residual*, similarly to photogrammetric bundle block adjustment. As a matter of fact, the experiments reported show that the method introduced in this paper achieves RMS error values (wrt. ground control points) practically equal to those obtained by photogrammetric bundle block adjustment.

## 2. Procrustes Analysis and Photogrammetry

Let us start this section by summarily presenting the main characteristics of the generalized (isotropic) Procrustes analysis in Photogrammetry and laser scanning applications. Afterwards, the anisotropic row-scaling Procrustes analysis will be presented, and its capabilities to successfully solve the exterior orientation of one image and the bundle block adjustment problem will be emphasized.

### 2.1. Registration of multiple 3-D models

As well known, photogrammetric relative orientation and laser scanning can provide numerical 3-D models of real objects by sampling the positions of a set of representative surface points. Depending on the extension and on the shape complexity of the geometric entity to be surveyed, its complete acquisition often leads to the creation of a set of partial and independent 3-D models. These parts must be joined together to reconstruct the complete object model into a unique frame by matching common points or features, or by directly aligning portions of corresponding surfaces.

The registration of multiple 3-D models or *n-view registration* problem requires to simultaneously transform into a unique mean coordinate system a set of  $m \geq 2$  models, each composed of  $n$  points coordinates in  $\mathbb{R}^3$  defined in a different reference frame.

If these models are expressed by  $m n \times 3$  matrices  $A_1, \dots, A_m$ , the problem is equivalent to:

$$\min \sum_{i < j}^m \|(\lambda_i A_i R_i + \mathbf{1} \mathbf{t}_i^T) - (\lambda_j A_j R_j + \mathbf{1} \mathbf{t}_j^T)\|_F^2 \quad (1)$$

---

<sup>1</sup>A projective/Euclidean model differs from the true one by a projectivity/similarity transformation.

where  $\mathbf{1}$  is the all-ones vector and  $(\lambda_i, R_i, \mathbf{t}_i)$  are the parameters of a similarity (a.k.a. Helmert) transformation. This is a generalized (isotropic) Procrustes analysis (GPA) model (Gower, 1975), whose solution allows to directly register all the 3-D models into a unique mean reference frame, minimizing a geometric error.

The analogy with the photogrammetric independent models block adjustment is evident:

- the number of the  $A_i$  matrices is equal to the number  $m$  of the models composing the block adjustment. The matrices contain the coordinates of the available points for each model;
- all matrices  $A_i$  have the same dimension, equal to the global number  $n$  of the block adjustment observations by the coordinate space dimension (usually 3);
- in the case of missing data, the generic matrix  $A_i$  has specified components only for the points belonging to the  $i$ -th model, the other ones being unspecified.

Figure 1 explains these positions.

Figure 1: Analogy between a block adjustment scheme and a GPA problem with missing data (from Beinat and Crosilla (2001)).

Different iterative procedures have been proposed in the literature to minimize the cost function (1). In particular we will follow the approach of Commandeur (1991), which is based on the following property:

$$\sum_{i < j}^m \|A_i^p - A_j^p\|_F^2 = m \sum_{i=1}^m \|A_i^p - S\|_F^2 \quad (2)$$

where  $A_i^p = \lambda_i A_i R_i + \mathbf{1} \mathbf{t}_i^\top$  is the matrix  $A_i$  transformed by the similarity with parameters  $\lambda_i, R_i, \mathbf{t}_i$ , and  $S = \frac{1}{m} \sum_{i=1}^m A_i^p$  is the *centroid* of the group of matrices, or *mean shape*: its rows are the coordinates of the geometric centroid of corresponding transformed points.

By comparing formula (2) with the objective function (1) it is possible to define a solving criterion based on iterative computation of the centroid, that leads to great advantages in terms of simplicity and efficiency. In summary, these are the two steps that are repeated until stabilization of the centroid  $S$ :

- a) The solution of independent similarity transformations for each matrix  $A_i^p$  with respect to the centroid  $S$ ;

- b) The computation of the centroid  $S$  following the sequential updating of matrices  $A_i^p$ .

Step a) is a simple extended orthogonal Procrustes analysis (EOPA) model, whose solution will be formulated ahead in this section.

In case of missing data, the problem can be handled with the procedure presented in Crosilla and Beinat (2002), where for each matrix  $A_i$ , a diagonal matrix  $M_i$  can be inserted, containing unit values along the main diagonal in case the corresponding row of  $A_i$  is specified and zero on the contrary. Hence, the cost function can be written as:

$$\sum_{i < j}^m \text{tr} (A_i^p - S)^\top M_i (A_i^p - S) \quad (3)$$

where the centroid  $S$ , is now defined as:

$$S = \left( \sum_{k=1}^m M_k \right)^{-1} \sum_{i=1}^m M_i A_i^p \quad (4)$$

### 2.1.1. Extended Orthogonal Procrustes Analysis

We outline here the extended orthogonal Procrustes analysis model; for more details please refer to Schönemann and Carroll (1970). The EOPA model considers the least squares similarity transformation between two matrices. Given matrices  $A$  and  $B$ , with the same meaning as  $A_i$  in the preceding paragraph, EOPA allows to define an orthogonal matrix  $R$ , a translation vector  $\mathbf{t}$  and a global scale factor  $\lambda$  which attain:

$$\min \|B - \lambda A R - \mathbf{1} \mathbf{t}^\top\|_F^2 \quad (5)$$

subject to the condition:  $R^\top R = R R^\top = I$ . The solution is the following. Let

$$Y = A^\top (I - \mathbf{1} \mathbf{1}^\top / n) B; \quad (6)$$

perform the Singular Value Decomposition of  $Y$

$$Y = U D V^\top$$

and set

$$R = U \text{diag}(1, 1, \det(UV^\top)) V^\top. \quad (7)$$

The central diagonal matrix is necessary if  $R$  is required to be a rotation matrix (Wahba, 1965; Arun et al., 1987).

Once  $R$  is known, the scale factor  $\lambda$  can be computed as:

$$\lambda = \frac{\text{tr}(R^\top A^\top (I - \mathbf{1} \mathbf{1}^\top / n) B)}{\text{tr}(A^\top (I - \mathbf{1} \mathbf{1}^\top / n) A)}. \quad (8)$$

And, finally:

$$\mathbf{t} = (B - \lambda A R)^\top \frac{\mathbf{1}}{n}. \quad (9)$$

A proof can be derived with the Lagrangian method as in Schönemann and Carroll (1970) (the AEOPA solution reported in Appendix A works along the same line). Related methods that aim at estimating the rotation can be found in (Wahba, 1965) and (Arun et al., 1987).

Summarizing, the PROCRUSTEAN MULTIPLE MODELS REGISTRATION (PMMR) algorithm is the following<sup>2</sup>:

---

### Algorithm 1 PMMR

---

**Input:** a set of 3-D models  $A_i \quad i = 1 \dots m$

**Output:** scale  $\lambda_i$ , translation  $\mathbf{t}_i$  and attitude  $R_i$  of each model

1. Initialize  $A_i^p = A_i \quad \forall i$
  2. Compute centroid  $S = \frac{1}{m} \sum_{i=1}^m A_i^p$
  3. Register each model  $A_i^p$  to  $S$ :
    - (a) Compute  $R_i = U \text{diag}(1, \text{det}(UV^\top)) V^\top$  with  $UDV^\top = A_i^{p\top} (I - \mathbf{1} \mathbf{1}^\top/n) S$ ;
    - (b) Compute  $\lambda_i = \frac{\text{tr}(R_i^{p\top} (I - \mathbf{1} \mathbf{1}^\top/n) S)}{\text{tr}(A_i^{p\top} (I - \mathbf{1} \mathbf{1}^\top/n) A_i^p)}$ ;
    - (c) Compute  $\mathbf{t}_i = (S - \lambda_i A_i^p R_i)^\top \mathbf{1}/n$ ;
    - (d) Update  $A_i^p = \lambda_i A_i R_i + \mathbf{t}_i^\top$
  4. Iterate from step 2 until convergence.
- 

The algorithm always converges (Commandeur, 1991), though not necessarily to the global minimum.

In the field of terrestrial laser scanning, the registration of multiple 3-D models via PMMR (with a rigid transformation) has been discussed in Beinat and Crosilla (2001); Toldo et al. (2010). In the field of photogrammetry, Crosilla and Beinat (2002) applied the PMMR to the solution of block adjustment by independent models. A good review can also be found in Akca (2003).

## 2.2. Exterior Orientation of one image

Given at least three control points and their projections, the *exterior orientation* problem requires to find a rotation matrix  $R$  and a vector  $\mathbf{c}$  (specifying attitude and position of the camera) such that:

$$\mathbf{p}_i = \zeta_i^{-1} R(\mathbf{s}_i - \mathbf{c}) \quad (10)$$

for some positive scalar  $\zeta_i$ , where

$\mathbf{s}_i$  is the coordinate vector of the  $i$ -th control point in the external system;

$\mathbf{c}$  is the coordinate vector of the projection center in the external system;

$\zeta_i$  is a positive scalar proportional to the “depth” of the point, i.e., the distance from the  $i$ -th control point to the plane containing the projection center and parallel to the image plane;

$R$  is the rotation matrix transforming from the external system to the camera system;

---

<sup>2</sup>In order to reduce clutter, the formulae in the algorithm refer to the case without missing data.

$\mathbf{p}_i$  is the coordinate vector of the  $i$ -th control point in the camera system, where the third component is equal to  $-c$ , the principal distance or focal length.

Equation 10 describes the central perspective projection and it leads to the well-known *collinearity equations*, as it expresses the fact that  $\mathbf{p}_i$  and  $R(\mathbf{s}_i - \mathbf{c})$  are collinear.

In a recent paper (Garro et al., 2012) the image exterior orientation problem has been solved using anisotropic row-scaling Procrustes analysis. We shall briefly review here its derivation and then move forward to the bundle block adjustment.

Expressing (10) with respect to  $\mathbf{s}_i$  yields:

$$\mathbf{s}_i = \zeta_i R^\top \mathbf{p}_i + \mathbf{c}. \quad (11)$$

After transposing

$$\mathbf{s}_i^\top = \zeta_i \mathbf{p}_i^\top R + \mathbf{c}^\top \quad (12)$$

and extending to  $n$  control points  $\mathbf{s}_1 \dots \mathbf{s}_n$  it results:

$$\underbrace{\begin{bmatrix} \mathbf{s}_1^\top \\ \vdots \\ \mathbf{s}_n^\top \end{bmatrix}}_S = \underbrace{\begin{bmatrix} \zeta_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \zeta_n \end{bmatrix}}_Z \underbrace{\begin{bmatrix} \tilde{\mathbf{p}}_1^\top \\ \vdots \\ \tilde{\mathbf{p}}_n^\top \end{bmatrix}}_P R + \underbrace{\begin{bmatrix} \mathbf{c}^\top \\ \vdots \\ \mathbf{c}^\top \end{bmatrix}}_{\mathbf{1}\mathbf{c}^\top}. \quad (13)$$

In compact form, the above formula reads:

$$S = ZPR + \mathbf{1}\mathbf{c}^\top \quad (14)$$

where  $P$  is the matrix by rows of image point coordinates defined in the camera frame,  $S$  is the matrix by rows of point coordinates defined in the external system,  $Z$  is the diagonal (positive) depth matrix.

One can recognize an instance of the extended orthogonal Procrustes analysis model (Sec. 2.1.1) generalized by the fact that the isotropic scale factor  $\lambda$  is substituted by an anisotropic scaling characterized by a diagonal matrix  $Z$  of different scale values. Indeed, according to Gower and Dijksterhuis (2004), this can be defined as an *anisotropic* extended orthogonal Procrustes analysis problem with *row scaling*.

To obtain the least squares solution for the anisotropic row-scaling Procrustes analysis model (14), one has to define a Lagrangian function and set to zero the partial derivatives with respect to the unknowns  $R$ ,  $\mathbf{c}$  and the diagonal matrix  $Z$  (details can be found in Appendix A).

The results are:

$$R = UV^\top \text{ with } UDV^\top = P^\top Z (I - \mathbf{1} \mathbf{1}^\top/n) S \quad (15)$$

$$\mathbf{c} = (S - ZPR)^\top \mathbf{1}/n \quad (16)$$

$$Z = (PP^\top \odot I)^{-1} (PR(S^\top - \mathbf{c}\mathbf{1}^\top) \odot I) \quad (17)$$

where  $\odot$  is the Hadamard (or element-wise) product.

The reader can notice that whereas in the solution of the extended orthogonal Procrustes analysis problem one can recover first  $R$ , that does not depend on the other unknowns, then the isotropic scale factor  $\lambda$ , and finally  $\mathbf{t}$ , in the anisotropic case the unknowns are entangled in such a way that there is no direct solution available. Gower and Dijksterhuis (2004) suggest an iterative procedure where each variable is alternatively estimated while keeping the others fixed<sup>3</sup>. This scheme is called *block relaxation* (de Leeuw, 1994) or *alternating least squares* (Young et al., 1976).

The PROCRUSTEAN EXTERIOR ORIENTATION (PEO) algorithm can be sketched as follows:

---

**Algorithm 2 PEO**


---

**Input:** control points  $S$  and their image coordinates  $P$

**Output:** position  $\mathbf{c}$  and attitude  $R$  of the camera

1. Start with any  $Z > 0$
  2. Compute  $R = U \text{diag}(1, 1, \det(UV^\top)) V^\top$   
with  $UDV^\top = P^\top Z (I - \mathbf{1} \mathbf{1}^\top/n) S$
  3. Compute  $\mathbf{c} = (S - ZPR)^\top \mathbf{1}/n$
  4. Compute  $Z = (PP^\top \odot I)^{-1} (PR(S^\top - \mathbf{c} \mathbf{1}^\top) \odot I)$
  5. Iterate from step 2 until convergence.
- 

It is understood that in step 4 one has to take care of negative values of  $Z$ , either by clipping or by non-negative least squares, as detailed in Appendix A.

Please note that while an iterative solution with guaranteed convergence was described by Bennani Dosse and Ten Berge (2010) for the cases of pre- and post-scaling of the columns (or “dimensions”), no analogous result is known in the literature for rows (or “configurations”) scaling, which is the problem we are dealing with here.

Solving exterior orientation with Algorithm 2 can be seen as iterating between two stages, namely:

- assuming known  $Z$ , apply the extended orthogonal Procrustes analysis to find attitude  $R$  and position  $\mathbf{c}$  of the image;
- given  $R$  and  $\mathbf{c}$ , solve for  $Z$  by finding the position along the (fixed) optical ray that minimizes the distance to the (known) 3-D points.

The solution of the anisotropic row-scaling Procrustes analysis problem finds  $Z, R$  and  $\mathbf{c}$  in such a way to minimize the sum of the squared norm of the difference vectors between 3-D reference points ( $S$ ) and the back-projected 2D points ( $P$ ) based on their estimated depths ( $Z$ ) and the estimated image attitude and position ( $R, \mathbf{c}$ ) (See Fig. 2).

In order to gain some understanding about the error being minimized – which is  $\|S - ZPR - \mathbf{1}\mathbf{c}^\top\|_F^2$  – let us

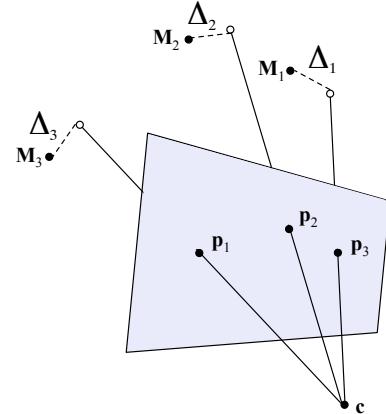


Figure 2: The estimated depth defines a 3-D point back-projected along the optical ray of the image point  $\mathbf{p}_i$ . The segment (perpendicular to the optical ray) joining this point and the corresponding reference 3-D point  $\mathbf{M}_j$  is the residual  $\Delta_j$ . The position and attitude of the camera plus the depth of the points are estimated in such a way to minimize the length of these  $\Delta_j$  for all the points, in a least squares sense.

rewrite:

$$S - ZPR - \mathbf{1}\mathbf{c}^\top = Z \underbrace{(Z^{-1}(S - \mathbf{1}\mathbf{c}^\top)R^\top - P)}_B. \quad (18)$$

The term  $(S - \mathbf{1}\mathbf{c}^\top)R^\top$  in the expression of the matrix  $B$  represents the 3-D coordinates of control points in the camera reference system; let  $[x_j, y_j, z_j]$  be the  $j$ -th row of such matrix and let  $[u_j, v_j, -c]$  be the  $i$ -th row of  $P$ . Then we have:

$$\begin{aligned} \|B\|_F^2 &= \|Z^{-1}(S - \mathbf{1}\mathbf{c}^\top)R^\top - P\|_F^2 = \\ &\sum_{j=1}^n \left\| \frac{1}{z_j} [x_j, y_j, z_j] - [u_j, v_j, -c] \right\|^2 = \\ &n(1 - c)^2 + \sum_{j=i}^n \left\| \left[ \frac{x_j}{z_j}, \frac{y_j}{z_j} \right] - [u_j, v_j] \right\|^2. \end{aligned} \quad (19)$$

The first term is constant while the second term is the sum of the squared distance between the measured image points  $[u_j, v_j]$  and the projection of the corresponding reference points. Hence minimizing  $\|B\|_F^2$  is equivalent to minimizing the sum of squared image coordinate residuals of the collinearity equations (i.e., the classical bundle adjustment error function).

Our method instead minimizes  $\|ZB\|_F^2$ , which is related to  $\|B\|_F^2$  by the following inequality:

$$\min(Z^2) \|B\|_F^2 \leq \|ZB\|_F^2 \leq \max(Z^2) \|B\|_F^2. \quad (20)$$

Since the entries of  $Z$  are bounded by physical constraints, i.e.,  $z_{\min} < z_i < z_{\max}$  with suitable constants  $z_{\min}$  and  $z_{\max}$ , when  $\|ZB\|_F^2$  approaches zero so does  $\|B\|_F^2$  and viceversa.

The insight to be gained from this application is that whereas long-established orthogonal Procrustes analysis

<sup>3</sup>It would be sufficient to remove the translation, though, to obtain a simpler problem with closed form solution (Gower and Dijksterhuis, 2004)

seeks transformations in 3-D space, the anisotropic extension allows to deal with problems involving a projection to 2D space by introducing an auxiliary unknown, the depth of the points, which allows to back-project 2D points into 3-D space, thereby restoring the 3-D problem.

### 3. Procrustean Bundle block adjustment

Let us now consider  $m$  images depicting the same  $n$  3-D tie-points  $\mathbf{s}_1 \dots \mathbf{s}_n$ . In the *bundle block adjustment* problem it is required to simultaneously find the image exterior orientation parameters and the tie-points 3-D coordinates that minimize a geometric error, without introducing intermediate models, as opposed to the block adjustment by independent models. Building on the previous section, we derive here a novel procrustean solution to the bundle block adjustment problem.

Starting from Eq. (14), which is an equivalent formulation of the collinearity equations, we can write for each image  $i$ :

$$S = Z_i P_i R_i + \mathbf{1} \mathbf{c}_i^\top. \quad (21)$$

In this formula the image coordinates  $P_i$  are known, but all the other quantities are unknown, including the 3-D points  $S$  and their depths  $Z_i$ . If  $Z_i$  were known, the problem would reduce to a GPA (with a rigid transformation), where the point sets to be aligned are the  $Z_i P_i \quad i = 1 \dots m$ .

If we rewrite Eq. (21) as  $P_i R_i = Z^{-1}(S - \mathbf{1} \mathbf{c}_i^\top)$  the model becomes the same as the STIMIDIO – modulo the inversion of the weight matrix – described at pg. 129 of Commandeur (1991), which in turn derives from the PINDIS (Lingoes and Borg, 1978) model. The iterative solution we propose is slightly different, though: along the same line as in the exterior orientation case, the proposal is to iterate between the following stages:

- assuming all the  $Z_i$  known, compute  $R_i, \mathbf{c}_i$  by solving a GPA problem (cfr. Eq. (1)), as in Algorithm 1:

$$\min \sum_{\ell < j} \| (Z_j P_j R_j + \mathbf{1} \mathbf{c}_j^\top) - (Z_\ell P_\ell R_\ell + \mathbf{1} \mathbf{c}_\ell^\top) \|_F^2;$$

- the putative 3-D points  $S$  are the centroids

$$S = \frac{1}{m} \sum_{i=1}^m (Z_i P_i R_i + \mathbf{1} \mathbf{c}_i^\top);$$

- given  $S$ , solve for  $Z_i$  independently in each image (as in Algorithm 2):

$$Z_i = (P_i P_i^\top \odot I)^{-1} (P_i R_i (S^\top - \mathbf{1} \mathbf{c}_i^\top) \odot I).$$

Since the global scale of the solution is discretionary, only the ratios of the  $Z_i$  are relevant. Therefore, at each iteration the  $Z_i$  are normalized to unit average, in order to better condition the convergence.

With respect to the classical photogrammetric bundle block adjustment, the error being minimized is different: Procrustean bundle block adjustment minimizes a geometric error in the 3-D space, whereas photogrammetric bundle block adjustment residuals are measured in the image plane. In particular, it is easy to see (Commandeur, 1991) that this procedure minimizes the following objective function:

$$\sum_{i=1}^m \| S - Z_i P_i R_i - \mathbf{1} \mathbf{c}_i^\top \|_F^2 \quad (22)$$

where each term of the sum represents the difference vector between 3-D tie-points ( $S$ ) and the back-projected 2D points ( $P_i$ ) based on their estimated depths ( $Z_i$ ) and the estimated image attitude and position ( $R_i, \mathbf{c}_i$ ) of image  $i$ . Let us call  $\Delta_{ij}$  each individual difference vector relative to exposure  $i = 1 \dots m$  and tie-point  $j = 1 \dots n$  (See Fig. 3). Overall, what is being minimized is the length of the residuals  $\Delta_{ij}$  for each exposure and each tie-point, in a least-squares sense.

The relationship with the classical photogrammetric bundle block adjustment cost function can be enlightened by following the same path as before (Eq. (18)-(20)). First we write the cost function Eq. (22) as  $\sum_i \| Z_i B_i \|_F^2$  where  $B_i = Z_i^{-1} (S - \mathbf{1} \mathbf{c}_i^\top) R_i^\top - P_i$ , and then prove that: i)  $\sum_i \| B_i \|_F^2$  is equivalent to the photogrammetric bundle block adjustment residual up to an additive constant ii) PBBA error is pinched between two functions proportional to the bundle adjustment error:

$$\begin{aligned} m \min_i \min(Z_i^2) \sum_{i=1}^m \| B_i \|_F^2 &\leq \sum_{i=1}^m \| Z_i B_i \|_F^2 \\ \sum_{i=1}^m \| Z_i B_i \|_F^2 &\leq m \max_i \max(Z_i^2) \sum_{i=1}^m \| B_i \|_F^2. \end{aligned} \quad (23)$$

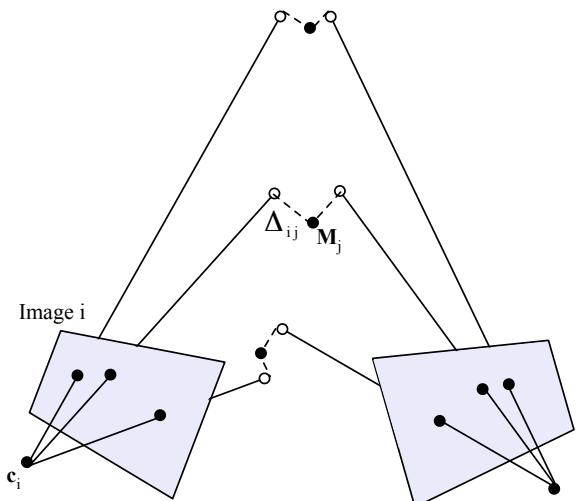


Figure 3: Consider the rays emanating from each point as “sticks” in space. PBBA optimize camera position and attitude and the length of each stick so as to bring the sticks’ endpoints as close as possible to each other, i.e., it minimizes the length of the  $\Delta_{ij}$ .

The PBBA solution corresponds to a free adjustment, as it does not involve any object space constraint. Points and images are expressed in an arbitrary reference system (inherent in the GPA), which can be brought into alignment with the desired coordinate frame via a similarity transformation, solving an absolute orientation problem with given ground control points.

The PROCRUSTEAN BUNDLE BLOCK ADJUSTMENT (PBBA) algorithm can be outlined as follows:

---

### Algorithm 3 PBBA

---

**Input:** 2D-2D correspondences  $P_i$  over  $m$  images  
**Output:** position  $\mathbf{c}$  and attitude  $R$  of the images, 3-D points  $S$

1.  $\forall i$  : Initialize  $Z_i > 0$
  2. Normalize the  $Z_i$  to unit average;
  3. Let  $A_i = Z_i P_i$ ; apply PMMR obtaining  $(R_i, \mathbf{c}_i)$  and centroid  $S$
  4.  $\forall i$  : Compute  $Z_i = (P_i P_i^T \odot I)^{-1} (P_i R_i (S^T - \mathbf{c}_i \mathbf{1}^T) \odot I)$
  5. Iterate from step 2 until convergence.
- 

If not all points are visible in all the images there are missing data, which are handled inherently by PMMR in step 2, while in step 3, for each image  $i$ , only the points visible in that image are considered.

The PBBA inherits from GPA its slow rate of convergence, which has been already pointed out in Akca (2003). Also Commandeur (1991) (pg. 134) comments on the slowness of his iterative solution of the STIMIDIO model. This trait has been partially cured in our implementation by introducing a damping factor greater than one in the updating of  $Z_i$ : the algorithm applies a correction to  $Z_i$  that is greater than the value computed according to step 4 of a factor ranging from 1 to 2, depending on the convergence rate.

As a final remark, although the nature of the two steps is completely different, one might see some resemblance of a customary resection-intersection cycle (Kraus, 1997, Sec. 4.1), as step 3 deals with the orientation while step 4 deals with the 3-D points. Also the iterative factorization methods described in the Introduction can be contrasted.

## 4. Experiments on simulated data

In this set of experiments, we tested PBBA on repeated trials with simulated data.

For each trial,  $n$  3-D tie-points have been randomly generated in a sphere of unit radius centered on the origin. Sixteen random cameras ( $m = 16$ ) looking toward the origin have been positioned in a  $60^\circ$  sector of the sphere, at an average distance of  $d$  units<sup>4</sup> from the origin. The

<sup>4</sup>In these experiments on simulated data, measures are in arbitrary “units”.

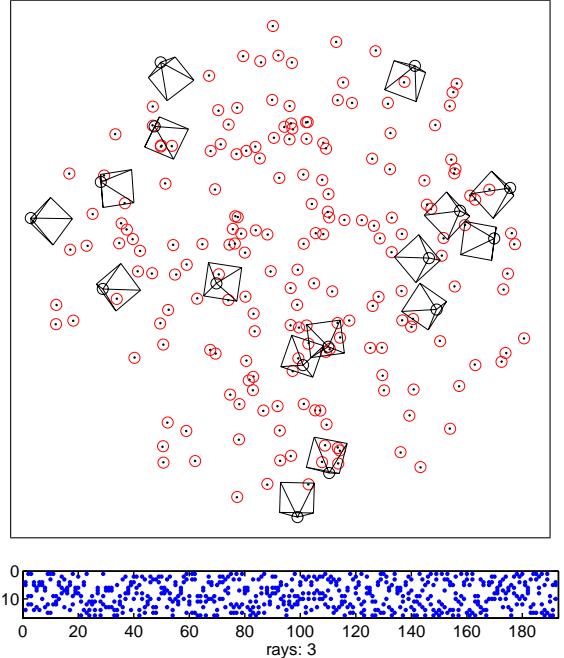


Figure 4: Top: A depiction of a simulated scene (camera stations and points) used in the experiment ( $n = 192$ ,  $d = 10$ ,  $p = 36$ ); Bottom: the simulated visibility matrix, showing which point (x-axis) is visible in which image (y-axis). Ray multiplicity is 3.

focal length has been chosen so as to yield a view angle of  $\{60^\circ, 120^\circ\}$  with an image size of  $1000 \times 1000$  pixels. Random noise has been added to the image coordinates. Missing points have been simulated by zeroing random elements in the visibility matrix.

A sample simulated scene is shown in Fig. 4. Figure 5 reports the convergence of error (as defined in Eq. (22)) and depth values for the first image. Please note that the initial value of  $Z$  needs not to have any relationship with the actual average depth of the 3-D points, for the global scale of the solution is meaningless.

Figure 6 shows the average and standard deviation of the residuals  $\Delta_{ij}$  for each point, over the number  $m$  of exposures. These residuals are, for each tie-point  $j$  imaged in exposure  $i$ , the distance from the back-projected point (according to its estimated depth) to its centroid, which represents the estimate of the actual 3-D point (see Fig. 3).

In the experiment aimed at assessing (empirically) the accuracy of the method, the parameters were set to:  $n = 96$ ,  $d = 10$ ,  $p = 36$  ( $p$  is the number of tie-points visible in each image). Increasing random noise with standard deviation  $\sigma = \{0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5\}$  has been added to image points coordinates and 25 trials for each noise level have been averaged. In each trial first PBBA is run and then photogrammetric bundle block adjustment, starting from the output of the former. The output of both methods have been compared after a least-squares alignment with the ground-truth 3-D points (Fig. 7 - left). The RMS error for PBBA is only slightly higher than the error achieved by photogrammetric bundle block adjustment; it

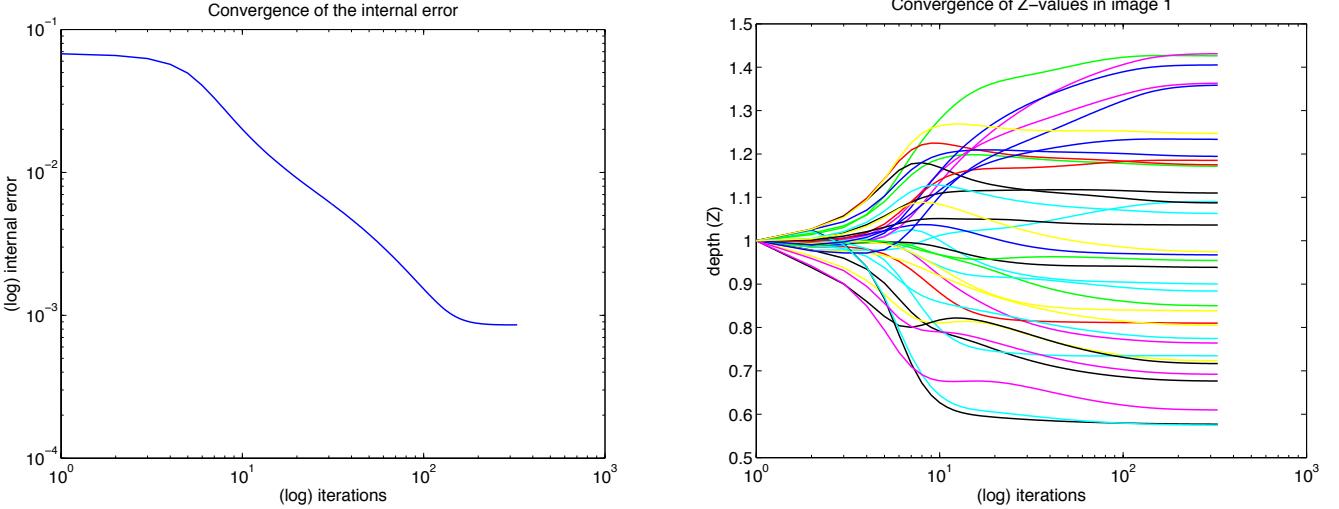


Figure 5: Convergence of error (left) and depth values (right) for the simulated scene of Fig. 4. Please note that some axes are logarithmic.

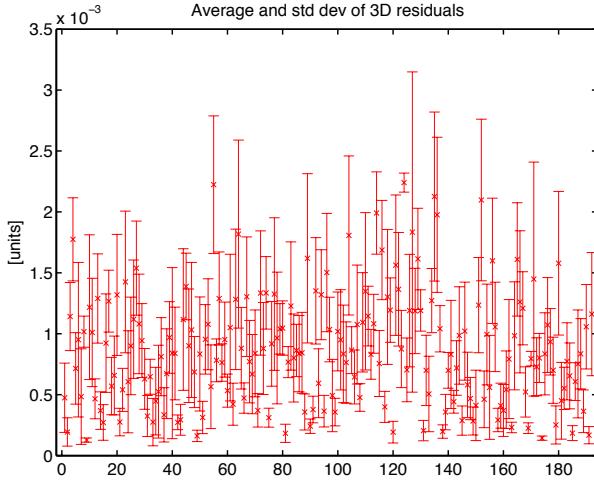


Figure 6: Average and standard deviation of the  $\Delta_{ij}$  (defined in the text) for the simulated scene of Fig. 4.

can be considered equal for all practical purposes. Note how the RMS error of 3-D points position grows linearly as the standard deviation of the noise increases. Figure 7 - right reports the root of reference variance of the image coordinate residuals for the same experiments.

A further experiment was designed for characterizing the failure rate, i.e., counting how many times PBBA failed to reach the correct solution in 100 random trials, starting with the usual uninformed initialization ( $Z = 1$ ). A random noise with 1 pixel standard deviation has been added to image coordinates. The parameters considered in this simulation are the distance  $d$  of the camera from the origin and the number  $p$  of tie-points visible in each image.

For each value of  $d = \{2, 10, 20\}$  3-D tie-points have been stretched in the  $XY$  plane so as to fit in the view frustum, while the  $Z$  range is kept constant (to the original two units), so as to give rise to increasing values for the

distance to  $Z$ -range ratio, namely  $\{1, 5, 10\}$  respectively.

The number of visible tie-points per image has been set to  $p = \{18, 36, 54\}$ . When the number of images and the number of visible tie-points per image are fixed, total number of points and the ray multiplicity become dependent. In one case the total number of points has been kept fixed to  $n = 96$  and the ray multiplicity took the values  $\{3, 6, 9\}$ . In another case ray multiplicity has been fixed to 3 and the total number of points took values  $n = \{96, 192, 288\}$ .

Figure 8 reports the results. When the ray multiplicity is greater than 3 (top row,  $p = \{36, 54\}$ ) the convergence rate is 100%. The case of ray multiplicity= 3 is analyzed in the bottom row. The case  $p = 18$ , view angle of  $60^\circ$  (with  $d = 2$ , in particular) is the worst; convergence rate is at least 95% in all the other cases.

Figure 9 reports the median RMS error on 3-D points position for this experiment. The RMS error is in percentage with respect to the radius of the point cloud (i.e., 1% is 0.01 units on 3-D points in a sphere with radius 1). This error is always below 2% and in most cases is below 1%. As expected, the error increases with the distance and decreases with the number of tie-points visible in each image. It also worsens when switching from a field of view of  $60^\circ$  (left column) to  $120^\circ$  (right column).

A final but important remark: although the convergence of PBBA is slow near the optimum, only few iterations are needed to dramatically reduce the error and obtain approximately correct depths (see Fig. 5 which is a paradigmatic example). This might suggest a combined use of PBBA and photogrammetric bundle block adjustment: PBBA is run for some iterations ( $\approx 20$ ) and then photogrammetric bundle block adjustment takes on and quickly converges to the minimum.

As for the computational cost, one iteration of photogrammetric bundle block adjustment requires approximately the same time spent in 7.4 iterations of PBBA, considering a MATLAB implementation of both running

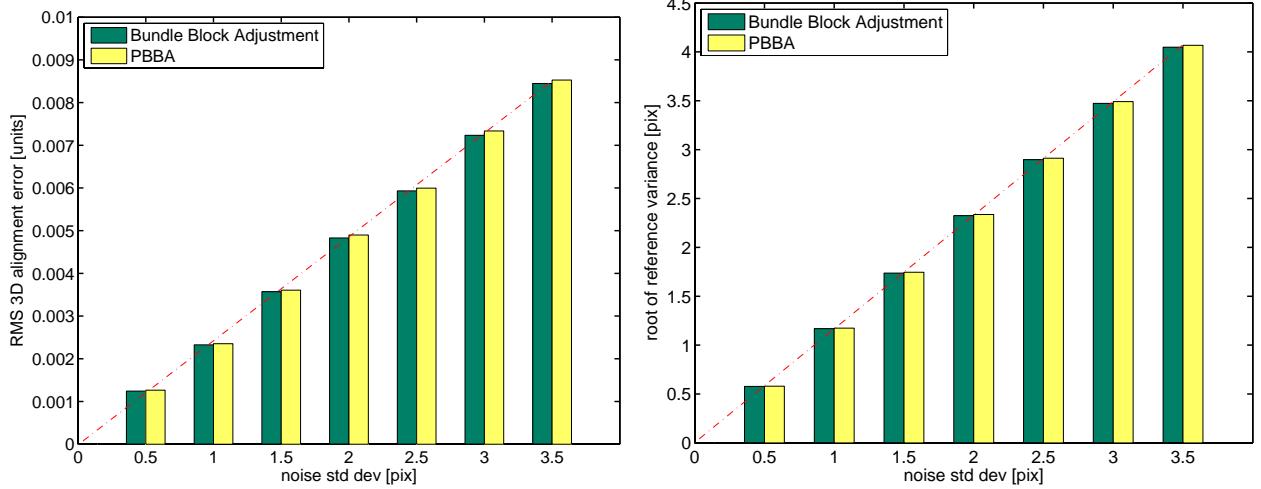


Figure 7: RMS error on 3-D point position (left) and root of reference variance of the image coordinate residuals (right) vs standard deviation of random noise added to image tie-points. Dark green is the error obtained by photogrammetric bundle block adjustment, while yellow is the error achieved by PBBA. The regression line is also plotted (in red).

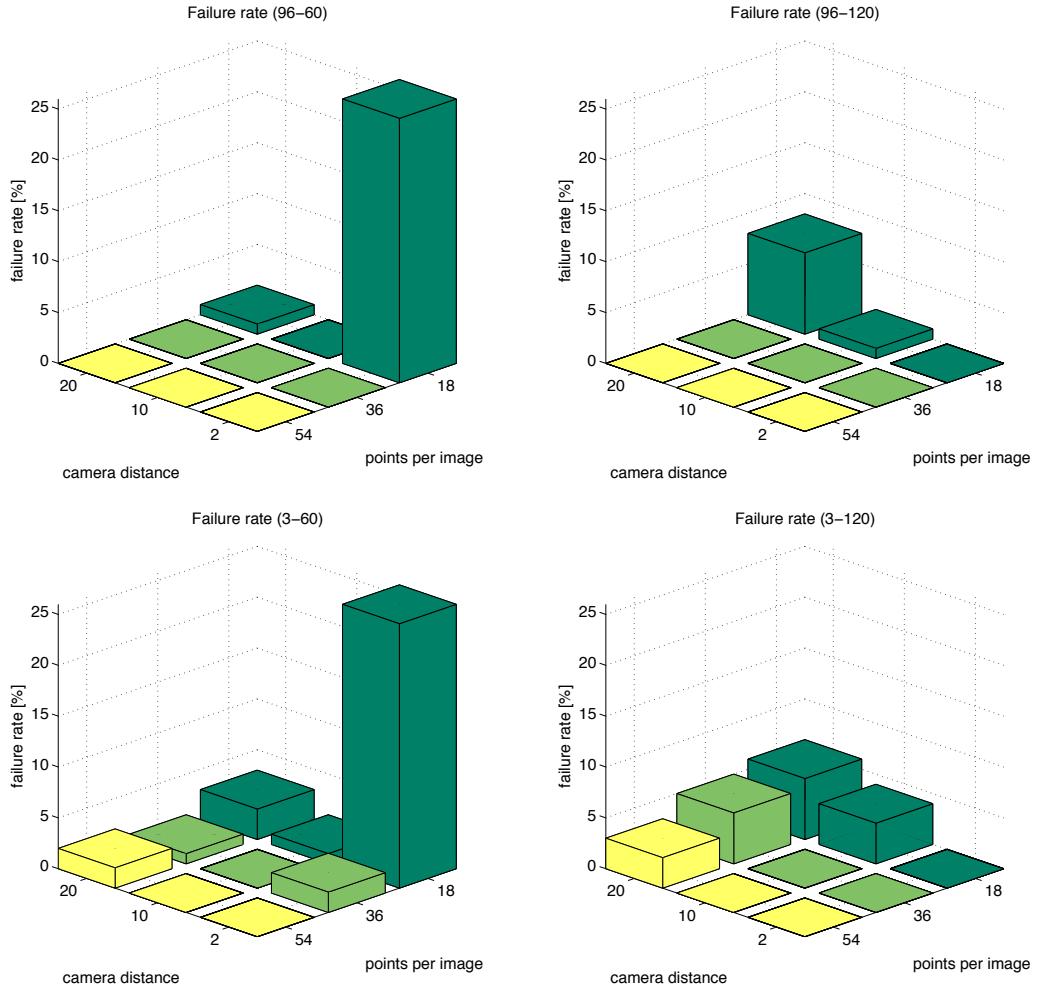


Figure 8: Failure rate of PBBA as a function of camera distance  $d$  and number of tie-points visible in each image. Top row: Corresponding ray multiplicity values are [3, 6, 9], where  $n = 96$ . Bottom row: Corresponding total number of points are [96, 192, 288], where the ray multiplicity is fixed to 3. Left/right column correspond to 60°/120° field of view, respectively.

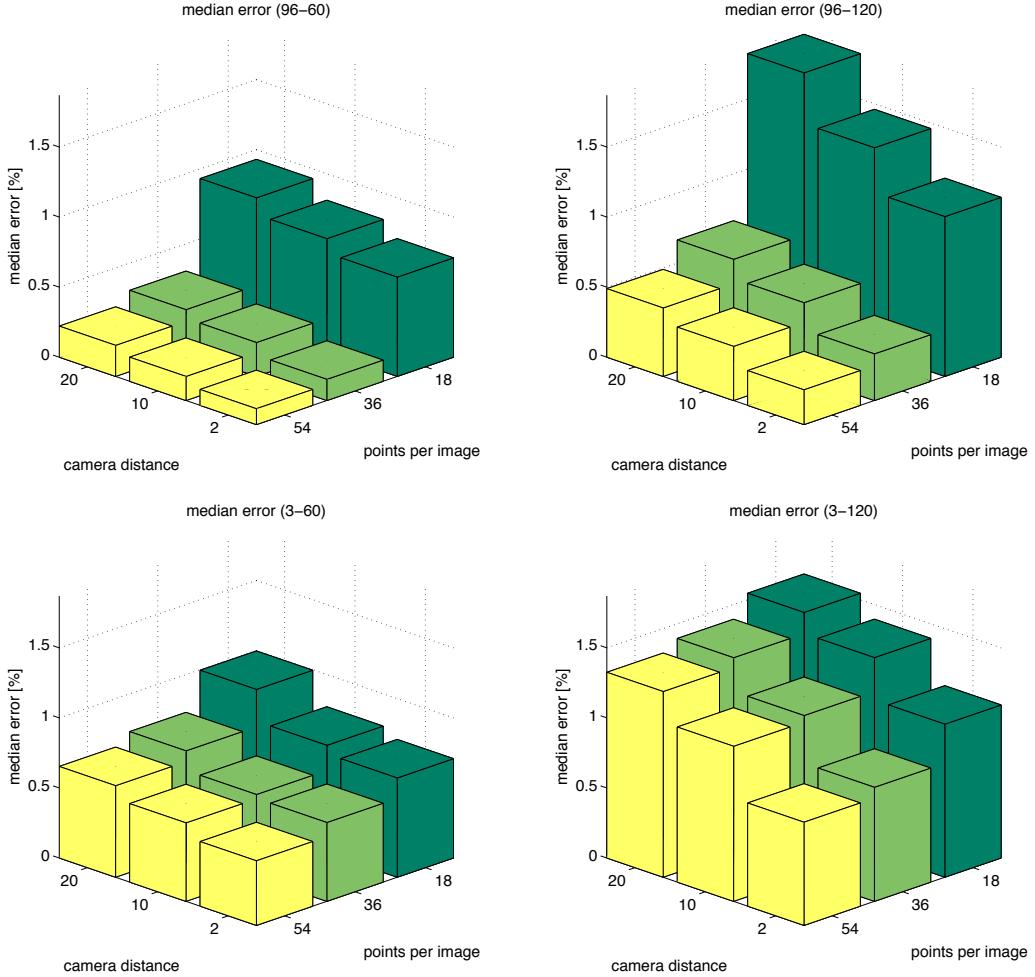


Figure 9: Median RMS error of PBBA as a function of camera distance  $d$  and number of tie-points visible in each image  $p$ . The error is in percentage with reference to points distributed on a unit sphere. Top row: Corresponding ray multiplicity values are  $[3, 6, 9]$ , where  $n = 96$ . Bottom row: Corresponding total number of points are  $[96, 192, 288]$ , where the ray multiplicity is fixed to 3. Left/right column correspond to  $60^\circ/120^\circ$  field of view, respectively.

on the simulated data of Fig. 4.

## 5. Experiments on real data

In this section we report two experiments to better evaluate the algorithm in real applications. In particular we selected one close-range experiment and one aerial (UAV) application (Fig. 10). Table 1 gives the main features of these datasets. In both cases image tie-points have been automatically computed using the structure-from-motion pipeline proposed by Farenzena et al. (2009); Gherardi et al. (2010)<sup>5</sup>. The Matlab implementation of bundle block adjustment has been taken from the Vision Lab Geometry Library<sup>6</sup>.

The close-range data is described in Strecha et al. (2008). In particular we used the Herz-Jesu-P25 set, composed by

Table 1: Datasets features. Image size and cardinality, average depth, ground sampling distance (GSD).

	size	avg. Z	GSD.
Herz-Jesu-P25	3072 x 2048 x 25	3.4 m	5.2 mm
Hessighein	4288 x 2848 x 15	242 m	61 mm

25 images, which have been carefully oriented in the referenced work, using a laser scan as the ground truth.

For the Herz-Jesu-P25 image-set (Fig. 11) the final reprojection RMS error is 0.54 pixels, whereas the PBBA error (average distance of backprojected points to their centroid) is 2.0 mm (Fig. 12). The RMS error with respect to the reference camera stations is 9.2 mm; Fig. 13 reports the error with respect to ground control points, both as distance in space and as coordinates differences. After running photogrammetric bundle block adjustment, this error improves only slightly, reaching 8.6 mm; this is not relevant though, for both measures are well within the un-

<sup>5</sup>code available from <http://samantha.3dflow.net>

<sup>6</sup>code available from <http://vision.ucla.edu/vlg/>



Figure 10: Left: one sample image of the Herz-Jesu-P25 set; Right: one sample image of the near-vertical Hessigheim set.

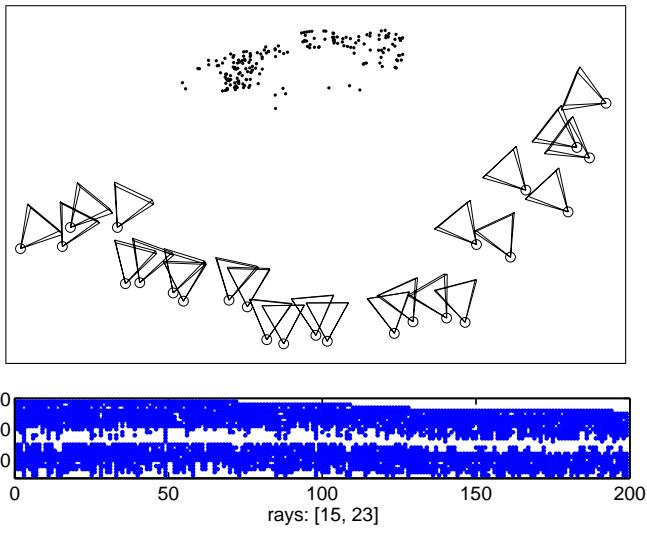


Figure 11: Top: A depiction of points and camera stations of the Herz-Jesu-P25 image-set; Bottom: the visibility matrix, showing which point (x-axis) is visible in which image (y-axis). The text below this matrix reports the minimum and maximum ray multiplicity.

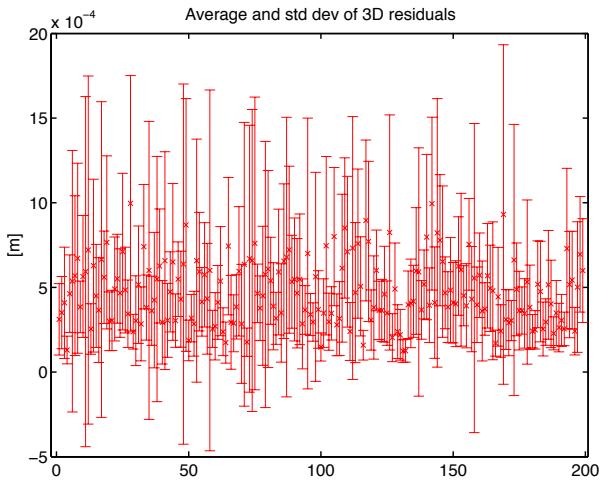


Figure 12: Average and standard deviation of  $\Delta_{ij}$  (defined in the text) for the Herz-Jesu-P25 image-set.

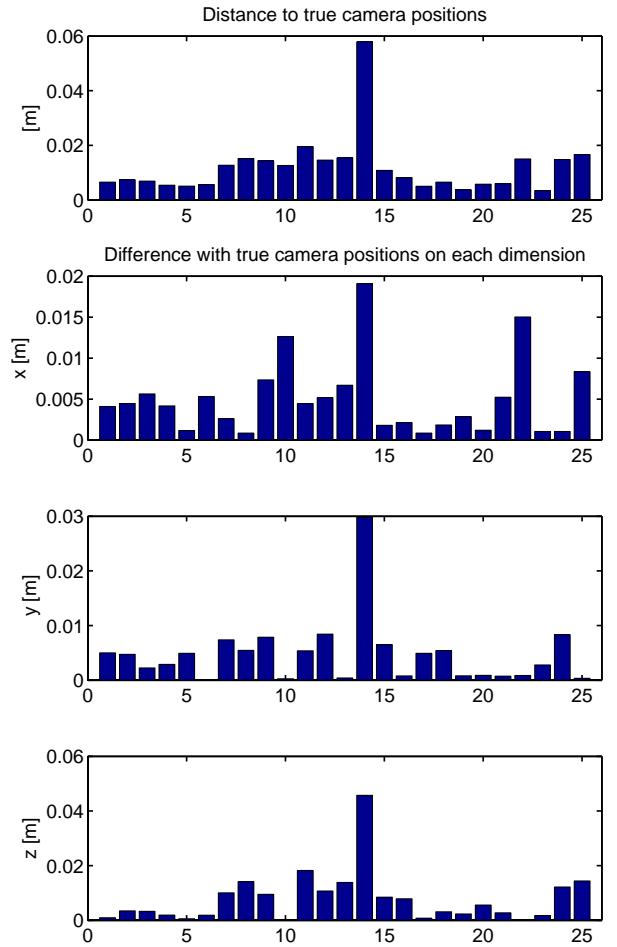


Figure 13: Top: Distance of computed to reference camera stations for the Herz-Jesu-P25 image-set. Bottom: Differences on each dimension (X-Y-Z).

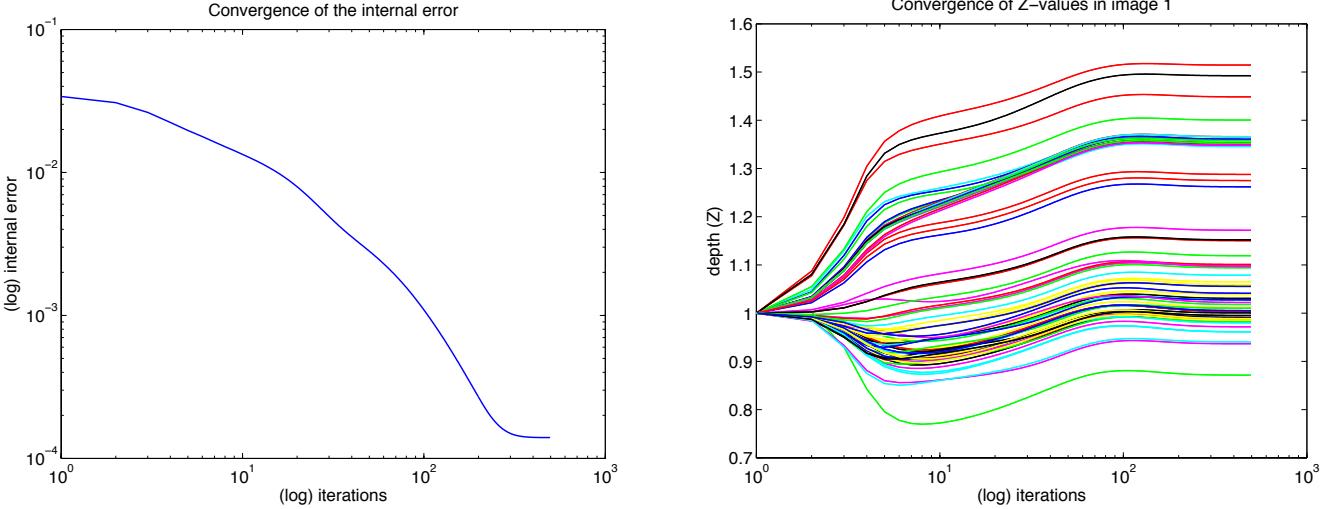


Figure 14: Convergence of error (left) and depth values (right) for the Herz-Jesu-P25 image-set. Please note that some axes are logarithmic.

certainty affecting the measured camera stations (reported in Strecha et al. (2008)). Fig. 14 shows the evolution of the error and depth values of points in the first image as the algorithm iterates.

The Hessigheim data is described in Cramer (2013). In particular we used only a small subset of those data, corresponding to 15 images. Interior orientation has been obtained by auto-calibration, as part of the structure from motion pipeline which provided the image tie-points.

For the Hessigheim image-set (Fig. 15) the final reprojection RMS error is 0.56 pixels, whereas the PBBA error (average distance of backprojected points to their centroid) is 29 mm (Fig. 16). The RMS error with respect to ground control points is 15.5 cm; Figure 17 reports the errors for each image and each dimension. After running photogrammetric bundle block adjustment, this error reaches 15.2 cm. This improvement however has little significance, if compared with the scale of the problem. Convergence can be appreciated in Fig. 18.

Please note that, in accordance with the remark on the global scale already made before, in both cases PBBA starts with  $Z = 1$  even if actual average depths are completely different (3.4m and 242m respectively).

The rundown of these experiments on real data is that the qualitative behavior of PBBA on simulated data is confirmed in all regards.

## 6. Conclusions

A new approach to bundle block adjustment based on anisotropic orthogonal Procrustes analysis has been presented. Beside its theoretical interest and its simplicity, this algorithm has the advantage of converging to the correct solution in most cases, with a zero-information initialization.

According to our empirical evaluation, PBBA and photogrammetric bundle block adjustment are somehow com-

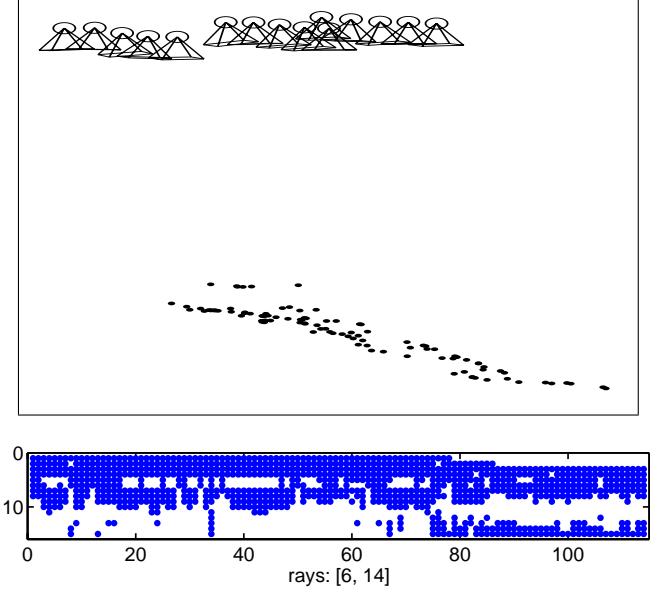


Figure 15: Top: A depiction of points and camera stations of the Hessigheim image-set; Bottom: the visibility matrix, showing which point (x-axis) is visible in which image (y-axis). The text below this matrix reports the minimum and maximum ray multiplicity.

plemental: while the former converges slowly in proximity of the solution but requires no specific initialization, the latter is faster but requires an educated guess of the solution to be started. Furthermore, photogrammetric bundle block adjustment yields theoretically optimal estimators – since it minimizes the observational error – whereas PBBA enlarges the range of convergence while minimizing a geometric error that is different but related to that of photogrammetric bundle block adjustment. In particular, we have proved that PBBA error is pinched between two functions proportional to the bundle adjustment error. Experiments show that there is no substantial difference in

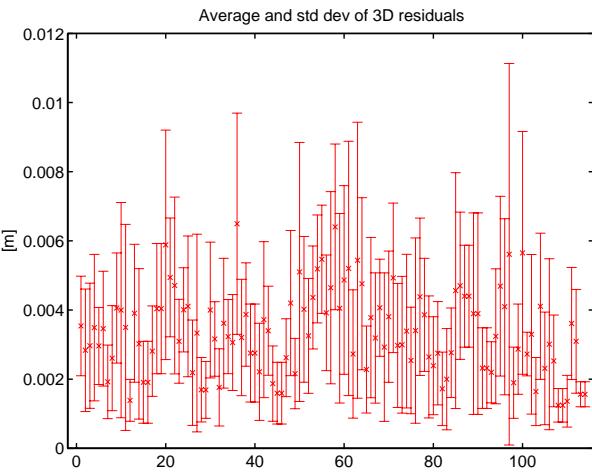


Figure 16: Average and standard deviation of the  $\Delta_{ij}$  (defined in the text) for the Hessigheim image-set.

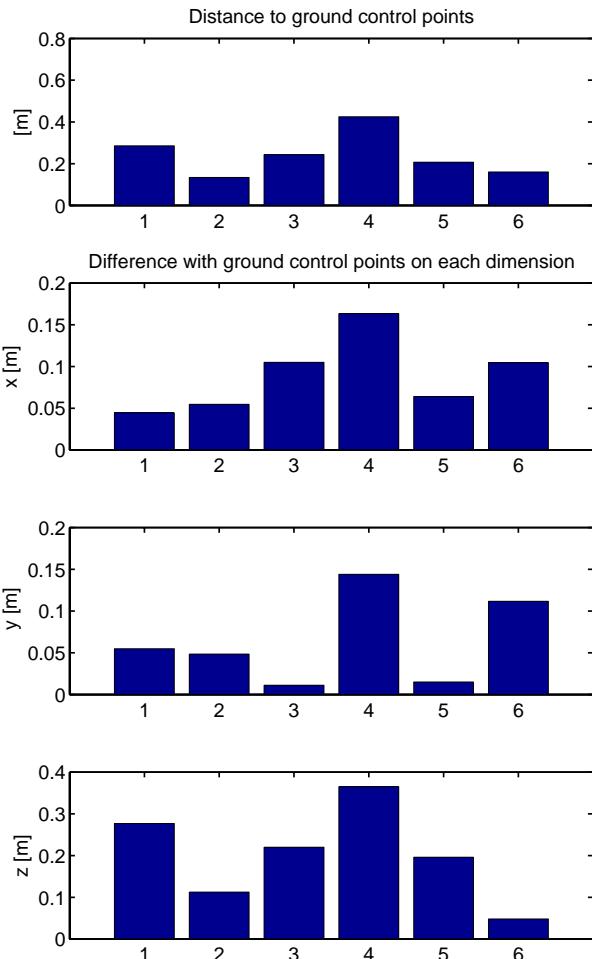


Figure 17: Top: Distance of computed 3-D points to ground control points for the Hessigheim image-set. Bottom: Differences on each dimension (X-Y-Z).

the RMS error of the results obtained by the two methods. A formal characterization of the accuracy of PBBA, however, would require the introduction of a stochastic model, which is left for future work. The route traced in (Fusiello et al., 2013) could be possibly followed to deal with the issue of sensitivity to outliers.

#### Acknowledgments

Thanks to N. Haala (ifp, Stuttgart) for providing the Hessigheim data, to R. Toldo (3Dflow s.r.l.) for running the structure from motion pipeline and providing tie-points, to M. Bennani Dosse for useful hints on the anisotropic row-scaling Procrustes analysis. The use of the dense multi-view stereo test images from EPFL (Strecha et al., 2008) and the UCLA Vision Lab Geometry Library is gratefully acknowledged.

#### Appendix A. Derivation of the anisotropic row-scaling Procrustes analysis solution

We shall report here the complete derivation of the anisotropic row-scaling Procrustes analysis solution, which is similar to the one reported by Schönemann and Carroll (1970) for the extended orthogonal Procrustes analysis.

To obtain the least squares solution for (14), let us make explicit the residual matrix  $\Delta$ :

$$\Delta = S - ZPR - \mathbf{1}\mathbf{c}^T. \quad (\text{A.1})$$

The problem is equivalent to the minimization of the Lagrangian function

$$F = \text{tr}(\Delta^T \Delta) + \text{tr}(L(R^T R - I)) \quad (\text{A.2})$$

where  $L$  is the matrix of Lagrangian multipliers. This can be solved by setting to zero the partial derivatives of  $F$  with respect to the unknowns  $R$ ,  $\mathbf{t}$  and the diagonal matrix  $Z$ .

Let us substitute (A.1) in (A.2):

$$\begin{aligned} F = & \text{tr}(S^T S) + \text{tr}(R^T P^T Z^T ZPR) + n \text{tr}(\mathbf{c}\mathbf{c}^T) - \\ & - 2\text{tr}(S^T \mathbf{1}\mathbf{c}^T) - 2\text{tr}(S^T ZPR) + \\ & + 2\text{tr}(R^T P^T Z^T \mathbf{1}\mathbf{c}^T) + \text{tr}(L(R^T R - I)) \end{aligned} \quad (\text{A.3})$$

The projection centre  $\mathbf{c}$  can be satisfied by equating to zero the partial derivative:

$$\frac{\partial F}{\partial \mathbf{c}} = 2n\mathbf{c} - 2S^T \mathbf{1} + 2R^T P^T Z^T \mathbf{1} = 0 \quad (\text{A.4})$$

Hence:

$$\mathbf{c} = (S - ZPR)^T \mathbf{1}/n \quad (\text{A.5})$$

Once the derivatives of  $F$  with respect to  $R$  and  $\mathbf{c}$  are set to zero, it results:

$$\begin{aligned} \frac{\partial F}{\partial R} = & P^T Z^T ZPR - P^T Z^T S + P^T Z^T \mathbf{1}\mathbf{c}^T + \\ & + R(L + L^T)/2 = 0 \end{aligned} \quad (\text{A.6})$$

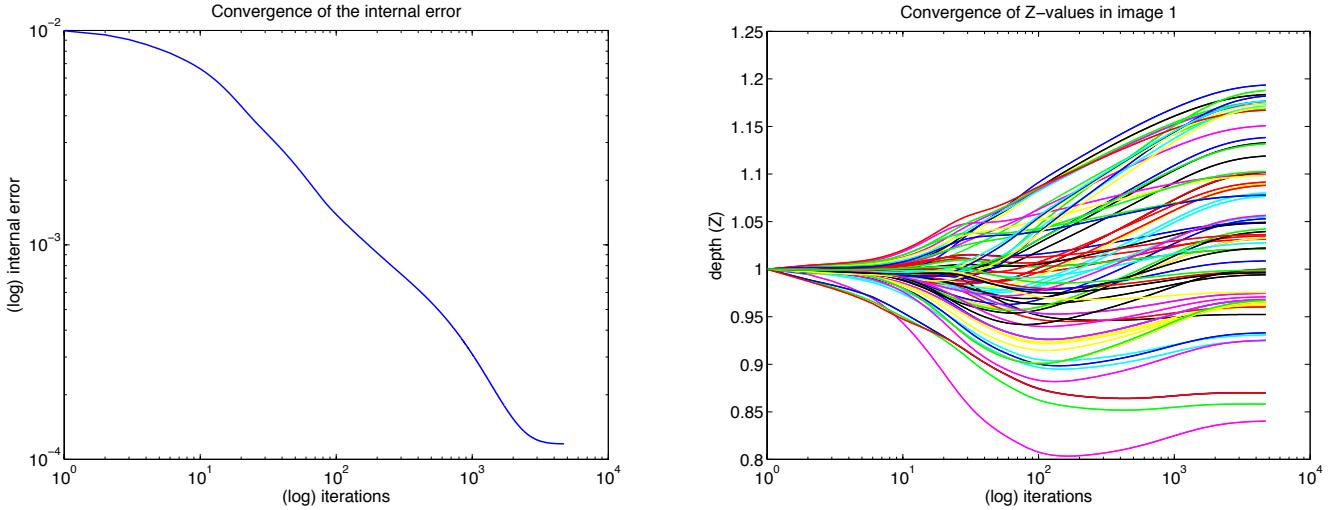


Figure 18: Convergence of error (left) and depth values (right) for the Hessigheim image-set. Please note that some axes are logarithmic.

where  $Q = (L + L^\top)/2$ .

Let us multiply (A.6) on the left by  $R^\top$ :

$$R^\top P^\top Z^\top ZPR - R^\top P^\top Z^\top S + R^\top P^\top Z^\top \mathbf{1}\mathbf{c}^\top + R^\top R(L + L^\top)/2 = 0 \quad (\text{A.7})$$

Since matrices  $R^\top P^\top Z^\top ZPR$  and  $(L + L^\top)/2$  are symmetric, then<sup>7</sup>

$$\text{sym}[R^\top P^\top Z^\top S - R^\top P^\top Z^\top \mathbf{1}\mathbf{c}^\top]. \quad (\text{A.8})$$

Substituting (A.5) in (A.8), it results

$$\text{sym}[R^\top P^\top Z^\top S - R^\top P^\top Z^\top (\mathbf{1}\mathbf{1}^\top/n)(S - ZPR)] \quad (\text{A.9})$$

which is equivalent to

$$\text{sym}[R^\top P^\top Z^\top S - R^\top P^\top Z^\top (\mathbf{1}\mathbf{1}^\top/n)S + R^\top P^\top Z^\top (\mathbf{1}\mathbf{1}^\top/n)ZPR] \quad (\text{A.10})$$

and finally:

$$\text{sym}[R^\top P^\top Z^\top (I - \mathbf{1}\mathbf{1}^\top/n)S + R^\top P^\top Z^\top (\mathbf{1}\mathbf{1}^\top/n)ZPR]. \quad (\text{A.11})$$

Since  $R^\top P^\top Z^\top (\mathbf{1}\mathbf{1}^\top/n)ZPR$  is symmetric, also the first term must be symmetric, i.e.,

$$\text{sym}[R^\top P^\top Z^\top (I - \mathbf{1}\mathbf{1}^\top/n)S]. \quad (\text{A.12})$$

Let us define the matrix  $T$  equal to

$$T = P^\top Z^\top (I - \mathbf{1}\mathbf{1}^\top/n)S \quad (\text{A.13})$$

<sup>7</sup>The predicate  $\text{sym}[\cdot]$  is true when the argument is a symmetric matrix.

Matrix  $R^\top T$  is symmetric, therefore the following condition must be satisfied

$$R^\top T = T^\top R \quad (\text{A.14})$$

that is equivalent to

$$TT^\top = RT^\top TR^\top \quad (\text{A.15})$$

Let  $T = UDV^\top$  be the SVD of  $T$ , with matrices  $V, U$  orthonormal. Substituting into (A.15) yields:

$$UD^2U^\top = RVD^2V^\top R^\top \quad (\text{A.16})$$

From (A.16)  $U = RV$  and finally  $R = UV^\top$ .

This formula only guarantees that  $R$  is orthogonal. The least squares estimate of a *rotation* matrix is obtained by (Wahba, 1965):

$$R = U \text{diag}(1, 1, \det(UV^\top)) V^\top. \quad (\text{A.17})$$

The least squares solution for the diagonal matrix  $Z$  can be obtained by setting to zero the partial derivatives of (A.3) with respect to  $Z$ .

$$\begin{aligned} \frac{\partial F}{\partial Z} &= \frac{\partial}{\partial Z} \text{tr}(R^\top P^\top Z^\top ZPR) - 2 \frac{\partial}{\partial Z} \text{tr}(S^\top ZPR) + \\ &\quad + 2 \frac{\partial}{\partial Z} \text{tr}(R^\top P^\top Z^\top \mathbf{1}\mathbf{c}^\top) \\ \frac{\partial F}{\partial Z} &= \frac{\partial}{\partial Z} \text{tr}(ZPRR^\top PZ^\top) - 2 \frac{\partial}{\partial Z} \text{tr}(PRS^\top Z) + \\ &\quad + 2 \frac{\partial}{\partial Z} \text{tr}(PR\mathbf{c}\mathbf{1}^\top Z) \end{aligned} \quad (\text{A.18})$$

Since  $Z$  is diagonal the derivative must be diagonal as well, so we introduce the Hadamard product  $X \odot I$  which is simply  $X$  with off-diagonal elements set to zero:

$$\begin{aligned} \frac{\partial F}{\partial Z} &= (Z(2PRR^\top P^\top) - 2PRS^\top + 2PR\mathbf{c}\mathbf{1}^\top) \odot I \\ \frac{\partial F}{\partial Z} &= (2ZPP^\top - 2PRS^\top + 2PR\mathbf{c}\mathbf{1}^\top) \odot I \end{aligned} \quad (\text{A.19})$$

By setting the derivatives to zero one obtains:

$$ZP\mathbf{P}^T \odot I = PR(S^T - \mathbf{c}\mathbf{1}^T) \odot I \quad (\text{A.20})$$

hence

$$Z = (PP^T \odot I)^{-1}(PR(S^T - \mathbf{c}\mathbf{1}^T) \odot I). \quad (\text{A.21})$$

The above formula may yield negative values for the diagonal entries of  $Z$ , which do not have a meaningful interpretation. When this happens one may clip them to zero or take a more principled approach and solve a constrained problem. Indeed, the last formula is equivalent to:

$$\text{diag } Z = (PP^T \odot I)^{-1} \text{diag}(PR(S^T - \mathbf{c}\mathbf{1}^T)). \quad (\text{A.22})$$

So, following Commandeur (1991), we can see  $\mathbf{v} = \text{diag } Z$  as the solution of the linear system of equations:

$$(PP^T \odot I)\mathbf{v} = \text{diag}(PR(S^T - \mathbf{c}\mathbf{1}^T))$$

which can be solved with non-negative least squares (Lawson and Hanson, 1974) if we require  $\mathbf{v} > 0$ .

## References

- Ackermann, E., 1962. Some results of an investigation into the theoretical precision of planimetric block adjustment. *Photogrammetria* 19 (8).
- Akca, D., 2003. Generalized procrustes analysis and its applications in photogrammetry. Internal Technical Report at the Institute of Geodesy and Photogrammetry - ETH, Zurich.
- Arie-Nachimson, M., Kovalsky, S. Z., Kemelmacher-Shlizerman, I., Singer, A., Basri, R., 2012. Global motion estimation from point matches. International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission.
- Arun, K. S., Huang, T. S., Blostein, S. D., September 1987. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9 (5), 698–700.
- Baarda, W., 1973. S-transformations and criterion matrices. Netherlands Geodetic Commission, Publications on Geodesy, New series, volume 5, number 1.
- Baetsle, P., 1956. Compensation des blocs photogrammetriques en altimetrie par relaxation. Bulletin de La Societe Belge de Photogrammetrie 4.
- Beinat, A., Crosilla, F., 2001. Generalized procrustes analysis for size and shape 3d object reconstruction. In: Gruen, Kahmen (Eds.), *Optical 3-D Measurement Techniques*. Wichmann Verlag, pp. 345–353.
- Bennani Dosse, M., Ten Berge, J., 2010. Anisotropic orthogonal procrustes analysis. *Journal of Classification* 27 (1), 111–128.
- Brand, M., 2002. Incremental singular value decomposition of uncertain data with missing values. In: *Proceedings of the European Conference on Computer Vision*. Springer, pp. 707–720.
- Brand, M., Antone, M., Teller, S., 2004. Spectral solution of large-scale extrinsic camera calibration as a graph embedding problem. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- Brown, D., 1958. A solution to the general problem of multiple station analytical stereotriangulation. Tech. Rep. AFMTC TR 58-8, Patrick Airforce Base, Florida.
- Brown, D., 1968. On first and second order partitioned regression. Internal working paper, DBA Systems Inc.
- Commandeur, J. J. F., 1991. Matching configurations. DSWO Press, Leiden.
- Cramer, M., 2013. The UAV@LGL BW project - a NMCA case study. In: Fritsch (Ed.), *Photogrammetric Week '13*. Wichmann, Berlin/Offenbach, pp. 165–179.
- Crosilla, F., Beinat, A., April 2002. Use of generalised procrustes analysis for the photogrammetric block adjustment by independent models. *ISPRS Journal of Photogrammetry & Remote Sensing* 56 (3), 195–209.
- Cunietti, M., 1968. Use of strips connected to blocks for large scale mapping. *Results of Experimental Research Organized by Commission B of the OEEPE from 1959 through 1966 - Frankfurt a.M.* de Leeuw, J., 1994. Block-relaxation algorithms in statistics. In: Bock, H. H., Lenski, W., Richter, M. M. (Eds.), *Information Systems and Data Analysis*. Springer-Verlag, pp. 308 – 325.
- Farenzena, M., Fusello, A., Gherardi, R., 2009. Structure-and-motion pipeline on a hierarchical cluster tree. In: *IEEE International Workshop on 3-D Digital Imaging and Modeling*. pp. 1489–1496.
- Forstner, W., 1985. The reliability of block triangulations. *Photogrammetric Engineering & Remote Sensing* 51 (8), 11371149.
- Fusello, A., Maset, E., Crosilla, F., 2013. Reliable exterior orientation by a robust anisotropic orthogonal procrustes algorithm. In: *Proceedings of the ISPRS Workshop 3D-ARCH 2013: 3D Virtual Reconstruction and Visualization of Complex Architectures*. Vol. XL-5/W1 of ISPRS Archives. Trento, Italy, pp. 81–86.
- Garro, V., Crosilla, F., Fusello, A., 2012. Solving the pnp problem with anisotropic orthogonal procrustes analysis. In: *Second Joint 3DIM/3DPVT Conference: 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT)*. pp. 262–269.
- Gherardi, R., Farenzena, M., Fusello, A., 2010. Improving the efficiency of hierarchical structure-and-motion. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 1594–1600.
- Gower, J., March 1975. Generalized procrustes analysis. *Psychometrika* 40 (1), 33–51.
- Gower, J. C., Dijksterhuis, G. B., January/Winter 2004. Procrustes problems. Vol. 30 of Oxford Statistical Science Series. Oxford University Press, Oxford, UK.
- Gruen, A., 1980. Precision and reliability aspects in close range photogrammetry. *Int. Arch. Photogrammetry* 11 (23B), 378391.
- Gruen, A. W., Baltsavias, E. P., 1986. Adaptive least squares correlation with geometrical constraints. In: *Computer Vision for Robots*. Vol. 0595 of *Proceedings of SPIE*. pp. 72–82.
- Gyer, M. S., 1967. The inversion of the normal equations of analytical aerotriangulation by the method of recursive partitioning. Tech. Rep. RADC-TR-67-69, Rome Air Development Center, NY.
- Hartley, R., Schaffalitzky, F., 2003. PowerFactorization: 3D reconstruction with missing or uncertain data. In: *Australia-Japan advanced workshop on computer vision*. Vol. 74. pp. 76–85.
- Hartley, R. I., Trampf, J., Dai, Y., Li, H., 2013. Rotation averaging. *International Journal of Computer Vision*.
- Heyden, A., 1997. Projective structure and motion from image sequences using subspace methods. In: *Scandinavian Conference on Image Analysis*. pp. 963–968.
- Jiang, N., Cui, Z., Tan, P., 2013. A global linear method for camera pose registration. In: *Proceedings of the International Conference on Computer Vision*.
- Kahl, F., Hartley, R., 2008. Multiple-view geometry under the  $l_\infty$ -norm. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30 (9), 1603–1617.
- Kaucic, R., Hartley, R. I., Dano, N. Y., 2001. Plane-based projective reconstruction. In: *Proceedings of the International Conference on Computer Vision*. pp. 420–427.
- Kennedy, R., Balzano, L., Wright, S. J., Taylor, C. J., 2013. Online algorithms for factorization-based structure from motion. CoRR abs/1309.6964. URL <http://arxiv.org/abs/1309.6964>
- Kraus, K., 1997. *Photogrammetry: Advanced methods and applications*. Vol. 2. Dümmler.
- Lawson, C. L., Hanson, R. J., 1974. Solving least squares problems. Prentice-Hall Series in Automatic Computation. Prentice-Hall, Englewood Cliffs.

- Lingoes, J., Borg, I., 1978. A direct approach to individual differences scaling using increasingly complex transformations. *Psychometrika* 43 (4), 491–519.
- Martinec, D., Pajdla, T., 2007. Robust rotation and translation estimation in multiview reconstruction. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
- Moulou, P., Monasse, P., Marlet, R., December 2013. Global Fusion of Relative Motions for Robust, Accurate and Scalable Structure from Motion. In: Proceedings of the International Conference on Computer Vision. Sydney, Australie.
- Oliensis, J., 1999. Fast and accurate self-calibration. In: Proceedings of the International Conference on Computer Vision.
- Oliensis, J., Hartley, R., 2007. Iterative extensions of the sturm/triggs algorithm: Convergence and nonconvergence. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29 (12), 2217–2233.
- Rother, C., Carlsson, S., Sep. 2002. Linear multi view reconstruction and camera recovery using a reference plane. *International Journal of Computer Vision* 49 (2-3), 117–141.
- Schönemann, P., Carroll, R., June 1970. Fitting one matrix to another under choice of a central dilation and a rigid motion. *Psychometrika* 35 (2), 245–255.
- Strecha, C., Von Hansen, W., Van Gool, L., Fua, P., Thoennessen, U., 2008. On benchmarking camera calibration and multi-view stereo for high resolution imagery. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–8.
- Sturm, P., Triggs, B., 1996. A factorization based algorithm for multi-image projective structure and motion. In: Proceedings of the European Conference on Computer Vision. Cambridge, UK, pp. 709–720.
- Toldo, R., Beinat, A., Crosilla, F., 2010. Global registration of multiple point clouds embedding the generalized procrustes analysis into an icp framework. In: Proceedings of the 5th International Symposium on 3D Data Processing, Visualization and Transmission.
- Triggs, B., McLauchlan, P., Hartley, R., Fitzgibbon, A., 2000. Bundle adjustment – a modern synthesis. In: Triggs, B., Zisserman, A., Szeliski, R. (Eds.), *Vision Algorithms: Theory and Practice*. Vol. 1883 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 298–372.
- Wahba, G., July 1965. A Least Squares Estimate of Satellite Attitude. *SIAM Review* 7 (3).
- Wilson, K., Snavely, N., 2014. Robust global translations with 1dsfm. In: Proceedings of the European Conference on Computer Vision. Vol. 8691 of Lecture Notes in Computer Science. Springer, pp. 61–75.
- Young, F., Leeuw, J., Takane, Y., 1976. Regression with qualitative and quantitative variables: An alternating least squares method with optimal scaling features. *Psychometrika* 41 (4), 505–529.
- Zhang, J., Boutin, M., Aliaga, D. G., 2006. Robust bundle adjustment for structure from motion. In: Proceedings of the International Conference on Image Processing. pp. 2185–2188.