

AUTOMATING THE PHOTOGRAMMETRIC BRIDGING BASED ON MMS IMAGE SEQUENCE PROCESSING

J. F. C. Silva ^{a*}, M. C. Lemes Neto ^b, V. Blasechi ^c

^a Unesp – Univ Est Paulista, FCT, Dept Cartografia, 19060-900 Presidente Prudente, SP, Brazil – (jfcsilva@fct.unesp.br)

^b UTFPR, Londrina, PR, Brazil – (mauricioneto@utfpr.edu.br)

^c Unesp – Univ Est Paulista, FCT, PPGCC, 19060-900 Presidente Prudente, SP, Brazil – (v.blasechi@gmail.com)

Commission I, ISPRS/IAG7

KEYWORDS: Photogrammetry, Triangulation, Orientation, Automation, Sequences, Terrestrial, Mobile, Mapping.

ABSTRACT:

The photogrammetric bridging or traverse is a special bundle block adjustment (BBA) for connecting a sequence of stereo-pairs and of determining the exterior orientation parameters (EOP). An object point must be imaged in more than one stereo-pair. In each stereo-pair the distance ratio between an object and its corresponding image point varies significantly. We propose to automate the photogrammetric bridging based on a fully automatic extraction of homologous points in stereo-pairs and on an arbitrary Cartesian datum to refer the EOP and tie points. The technique uses SIFT algorithm and the keypoint matching is given by similarity descriptors of each keypoint based on the smallest distance. All the matched points are used as tie points. The technique was applied initially to two pairs. The block formed by four images was treated by BBA. The process follows up to the end of the sequence and it is semiautomatic because each block is processed independently and the transition from one block to the next depends on the operator. Besides four image blocks (two pairs), we experimented other arrangements with block sizes of six, eight, and up to twenty images (respectively, three, four, five and up to ten bases). After the whole image pairs sequence had sequentially been adjusted in each experiment, a simultaneous BBA was run so to estimate the EOP set of each image. The results for classical (“normal case”) pairs were analyzed based on standard statistics regularly applied to phototriangulation, and they show figures to validate the process.

1. INTRODUCTION

Generally, a land-based mobile mapping system (MMS) has sensors to provide full orientation (position and attitude) for the stereo image pair sequence. Integrating positional and navigational data from satellite and inertial systems a solution to the image orientation problem is guaranteed (Tao, Li, 2007). While driving along streets and roads, the MMS acquires sequences of stereoscopic image pairs when two parallel cameras are pointed forward or backward the vehicle (Habib, 1998). With the outer orientation given by the sensors, the object points, along and aside the roads, selected in the image pairs, can be mapped by photogrammetric intersection (Silva et al., 2003).

Although such integration has been the standard method in MMS, a few researchers are developing a methodology entirely based only on image processing which means completely relied only on data extracted from the road image sequence to work out the image orientation problem in order to provide either an alternative or a complement to the satellite-inertial data integration system (Tao et al., 2001). In the realm of photogrammetry, we have proposed a technique called photogrammetric traverse (Silva et al., 1998) which is a special bundle block adjustment (BBA) for connecting a sequence of stereo-pairs and determining the exterior orientation parameters (EOP).

Of course, as no real external data are considered, we need to

provide an arbitrary datum such as a local Cartesian reference frame. We did it in a semi-automatic strategy based on optical flow to map conjugate points in the imagery domain along the sequence so they could be used as tie points (Barbosa et al., 2007).

We propose now to automate the photogrammetric bridging based on a fully automatic extraction of homologous points in stereo pair sequence to refer the EOP and tie points to the arbitrary object space Cartesian datum. The hypothesis is that we can have the orientation of the whole image sequence estimated with respect to the local reference system. In the following sections the detailed methodology, experiments, results, and analysis are presented and discussed in order to verify the alluded hypothesis.

2. BACKGROUND THEORY

In general both cameras and images require orientation with respect to a Cartesian reference frame (CRF).

2.1 Cartesian Reference Frames

2.1.1 CRF for Image Coordinates: An internal frame is defined inside the camera and it is used for referencing the interior orientation parameters (IOP), such as, focal length, principal point, lens distortion, and affinity parameters. Considering that the pair of cameras is mounted on a rigid metal structure so both optical axes are approximately horizontal and

* Corresponding author.

parallel to each other. In each camera the x-axis points to the right, the y-axis is upward, and the z-axis coincides with the optical axis pointing to the image sensor inside the camera, what defines the so called photogrammetric coordinate system (PCS) whose origin as usual coincides with the center of the image. Calibration procedures estimate the IOP of each camera and also the relative orientation parameters (ROP) between the pair of cameras.

2.1.2 CRF for Object Coordinates: The external orientation parameters (EOP) are referred in general to a CRF realized on the object. In this particular situation a local CRF is arbitrarily established by the camera pair system (CPS). Considering an observer behind the cameras, the CRF origin is set to the perspective center (PC) of the left camera; the stereo basis materializes the X-axis pointing to the PC of the right camera, Y-axis is upward, and Z-axis coincides with the optical axis pointing to the image sensor inside the camera or in other words points to the observer behind the CPS.

Both image and object Cartesian coordinates are then right-handed reference systems. Rotating angles between the corresponding image and object reference systems are positive counter clockwise. Both reference systems are related to each other by scale, translation, and rotation, and they are approximately parallel.

2.2 Image Navigation and Orientation.

Satellites for positioning constitute the currently called Global Navigation Satellite System (GNSS) that provides global coordinates and navigation parameters for many applications. Inertial sensors such as accelerometers and gyroscopes sense the specific force on a rigid body and its axis angular rates, respectively. Specific force and angular rate are suitably processed by the Kalman filter so we can estimate the position, velocity, and attitude (PVA) of the rigid body. Currently the integrated solution of satellite and inertial data provides an efficient method to orient the cameras and their products such as the image pairs (Rouzaud, Skaloud, 2011). For non-real time photogrammetric purposes, usually, the P is computed from satellite data and the A is from the inertial.

Not always we can have the combined satellite and inertial PVA estimates. Indoor images and dense urban street images, for instance, may not have the EOP computed from PVA as these estimates may fail due to the lack of satellite data and the inertial data accumulates noises because of high frequency rates, temperature, and the like.

Some computer vision and also mapping researchers started experiencing on techniques to orient images based only on image processing or combining it with other technologies. (Randeniya et al., 2008) presented the calibration results of an integrated visual and inertial system which is a passive technique suitable for indoor environment. Using only data from the visual system applied to an outdoor application, the authors obtained stimulating results for camera orientation. (Li and Sclaroff, 2007) published a solution that uses a stereo camera, optical flow, and correspondence to estimate the three dimensional movement fields for the objects in the scenes, and (Giachetti et al., 1998) for road navigation. (Griessbach et al., 2010) presented a multi sensor approach for realization of the orientation task in combination with an optimal fusion of the measurements that provides the best estimation accuracy of parameters. Depending on the applications the main inertial

sensors will be combined with one or more other specific sensors like optical sensor (stereo cameras), GPS and others. The main investigation points are the development of the methods and algorithms for complex state estimation and their implementation in a real-time software and hardware solution, which serves as a base navigation system for different applications, for example for indoor navigation, driver assistance for vehicles and trains or the estimation of exterior orientation for airborne and space borne camera systems and attitude control for small satellites. (Amami et al., 2014) produced a personal MMS based mainly on cameras and use of low cost GNSS and inertial sensors to provide a bundle adjustment solution with initial values. The system has been tested indoors and outdoors with different GPS coverage, surrounded features, and narrow and curvy paths. Tests showed that the system is able to work in such environments providing 3D coordinates of better than 10 cm accuracy.

(Silva et al., 2007) and (Barbosa et al., 2007) implemented a technique to orient the image pairs based only on image processing and photogrammetric techniques without any other sensor. The solution was based on the vehicle velocity estimated from the dense optical flow computation.

(Veth, 2011) synthesized the techniques and the progress related to navigation fully relied on images. He grouped the techniques in two classes of methods: one based on optical flow and the other on feature tracking. Both methods use the apparent movement of image regions among the sequential frames to determine the relative movement of a camera. Shortly every image based technique for positioning requires three basic operations: finding suitable regions in the images, corresponding regions to establish homologous points between them, and estimating the EOP. Each operation has its own drawbacks that are grouped in scientific and technologic problems such as, respectively, interest area selection, correspondence, and pose estimation. In photogrammetry, these operations are denominated: point or feature extraction, correspondence (correlation), and POE estimation.

2.3 Photogrammetric Bridging

(Roncella et al., 2005) used the term photogrammetric bridging. (Silva et al., 1998, 1999) had proposed photogrammetric traverse. It is a special BBA for connecting a sequence of terrestrial stereo-image (horizontal) pairs taken along the vehicle trajectory so that stereoscopic basis is approximately positioned perpendicular to the main surveying axis or that the optical axes are approximately parallel to the mentioned surveying axis. Of course, the BBA determines the exterior orientation parameters (EOP).

Figure 1 illustrates the concept where L_i, R_i represent the stereoscopic bases at instants i . Considering L for left and R for right, the positions where the images, i.e., the PC, are taken are represented by L_1 and R_1 at instant t_1 and so on for the sequential instants t_i . The object point P is situated between the third and the fourth bases and it appears in the images of bases 1, 2 and 3. Its spatial position (X, Y, Z) with respect to the local and arbitrary CRF (LACRF) can be estimated by simple intersection with third pair data or double intersection with second and third bases data or yet by triple intersection of the six images. By generalizing this principle we reach the concept of phototriangulation with the particularity that in such case the scale varies severely in each image. This is due to the range Y_p of the object point P to the bases varies from a few meters to

dozens ahead.

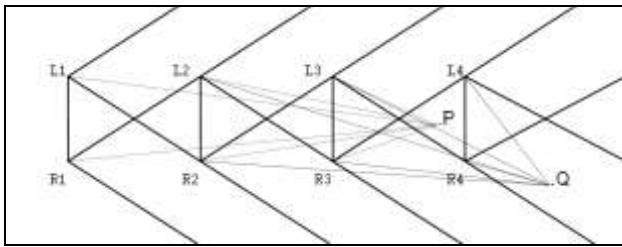


Figure 1 – Photogrammetric bridging or traverse – conceptual illustration (Silva, 1998)

2.4 SIFT Algorithm

The algorithm was introduced by (Lowe, 1999) and it is based on earlier contributions from many authors such as (Moravec, 1981) and (Brown; Lowe, 2002). The Scale Invariant Feature Transform (SIFT) algorithm is a set of four main steps that filters successively an image so it extracts features that are invariant to image scaling and rotation, and partially invariant to change in illumination and camera orientation. We suppose that when using images of a sequence the characteristics of invariance are kept especially in the high contrast regions of the images where it is possible to extract edges in general and points in particular.

The four steps were described by (Lowe, 2004): 1. Scale-space extrema detection which is the first stage of computation that searches over all scales and image locations. It uses a difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation; 2. Keypoint localization details a model to determine location and scale at each candidate location. Keypoints are selected based on measures of their stability; 3. The orientation assignment step attributes to each keypoint one or more orientation based on local image gradient directions. All future operations are performed on image data that has been transformed relatively to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations; 4. Keypoint descriptor closes the process when the local image gradients are measured at the selected scale in the region around each keypoint. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination.

In the realm of image orientation this process searches objects as points in the images. SIFT extracts a large amount of keypoints – the so called clouds of points – so that the consequence of an eventual error due to the scene variation is reduced. At the end of the process, the algorithm produces four attributes for each keypoint in a feature vector: x and y coordinates, magnitude, and orientation.

3. AUTOMATING THE PROCESS

The total sequence of the frontal images is processed sequentially by forming partial blocks of four, six, eight, or more images, that is to say partial blocks of two, three, four, or more image pairs as we will present in the coming sections. The photogrammetric-based orientation process demands automation.

It begins with the SIFT-based determination of a cloud of points from each image of a partial block. The correspondences among

cloud points are determined by a correlation technique. The image coordinates of all corresponding points are transformed to photo-coordinates. EOP are set arbitrarily to the first pair in order to establish a LACRF. The object space coordinates are approximated by photogrammetric intersection from EOP and photo-coordinates using the inverse collinearity equations. The block of four, six, eight, or more images are then phototriangulated. The process is repeated sequentially until the last base or pair is reached and all EOP of each image is estimated. This is called photogrammetric traverse or bridging.

The Matlab tools (v. 7.0) was used to implement the computational solution which is composed of SIFT algorithm, matching, intersection, and phototriangulation (BBA) besides using existing software and libraries to construct and visualize images and graphics.

3.1 Point Clouds by SIFT Algorithm

The relative position among objects in the scene is an important feature when we want to recognize objects. The locations are not expected to change significantly from an image to another one in the vicinity as it is the case of a sequence, otherwise many errors may occur. In this project the SIFT technique was applied to extract points with common characteristics among a block of images, instead of object recognition.

The amount of cloud points is directly proportional to the size of the image and to the threshold related to the difference between the first and second least distances considering the feature vectors. Consequently, so it is the computing time. Around 12 to 15 thousand of keypoints were found in each of the first four images of size 1920 x 1280 and around six to seven hundred in the 80% reduced size images (384 x 256). The images were reduced by bicubic interpolation.

3.2 Manifold Corresponding Points

Considering a block of four images (two pairs) such as L_1-R_1 form the pair 1, at epoch t_k , and L_2-R_2 form the pair 2, at epoch t_{k+1} . SIFT was applied to each image so it extracted four distinct point clouds. Taking the pair L_1-R_1 , the homologous points were found based on the minimum Euclidian distance of their feature vectors in the vicinity of every point of the left image (L_1), that is to say, the closest neighbor. Taking now the two left images (L_1 and L_2) of the first and second pairs, the homologous points of the first pair left image (L_1) were correlated with the points of the second pair left image (L_2) based on the same criteria resulting in a set of forward homologous points. The procedure was repeated for L_2-R_2 , and finally for R_1-R_2 , so the correspondence technique was applied twice for the lateral (stereo) and twice for forward images.

The higher the threshold the larger is the quantity of the correspondences among the keypoints. Higher threshold than 0.6 may cause false matchings. That happened when using 0.8 for the original size images that produced 1862 four-fold keypoint correspondences (only 68 for the 80% reduced images). The threshold of 0.3 gave 89 four-fold correspondences for the original size and only 5 matched points for the 80% reduced.

As the partial blocks of the whole sequence of frontal terrestrial images may have more than two bases, SIFT algorithm is applied to all images of every partial sequence. Accordingly to the procedure above, the homologous or corresponding points

are automatically determined by four, six, eight or ten-fold image matching process. These points are then treated as tie points whose planar image coordinates (r,c) of the manifold correlated points are transformed to PCS (f, x₀, y₀). All these selected points will be used as tie points in the partial block or sequence.

3.3 Tie Points Terrain Coordinates by Photogrammetric Intersection

In order to automate the photogrammetric bridging process, the tie points must have their approximate object space coordinates computed by photogrammetric intersection. Working out the collinearity equations we can write the linearized and modified pseudo-observation equations based on the arbitrary EOP and tie point photo-coordinates after some algebraic manipulation. A linear least squares adjustment estimates the corresponding XYZ tie point coordinates without any approximation by solving a system (eq 1) of 3 x 3 normal equations (Silva et al, 2003).

$$x_t = (A^T A)^{-1} (A^T l_b) \quad (1)$$

The vector x_t contains the three XYZ coordinates to be used as approximate values in the BBA of the partial sequence. The design matrix A will range from 8 x 3 (double intersection) up to 20 x 3 elements (five-fold intersection). The first two rows of the design matrix A and the pseudo-observation vector (l_b) are expressed as they follow, respectively:

$$\begin{pmatrix} (x-x_0) \cdot r_{31} + f \cdot r_{11} & (x-x_0) \cdot r_{32} + f \cdot r_{12} & (x-x_0) \cdot r_{33} + f \cdot r_{13} \\ (y-y_0) \cdot r_{31} + f \cdot r_{21} & (y-y_0) \cdot r_{32} + f \cdot r_{22} & (y-y_0) \cdot r_{33} + f \cdot r_{23} \end{pmatrix} \quad (2)$$

$$\begin{pmatrix} (x-x_0) K + f \cdot I \\ (y-y_0) K + f \cdot J \end{pmatrix} \quad (3)$$

where I = (r₁₁ X_c+r₁₂ Y_c+r₁₃ Z_c), J = (r₂₁ X_c+r₂₂ Y_c+r₂₃ Z_c), K = (r₃₁ X_c+r₃₂ Y_c+r₃₃ Z_c); (x,y,f) are the photocoordinates and the calibrated focal length, respectively; (x₀,y₀) are the principal point coordinates; {X_c,Y_c,Z_c} are the object space PC coordinates; and r_{ij} the elements of the Eulerian rotation matrix.

Of course the row dimensions of A and l_b will vary from 8 to 20 accordingly to the four-fold up to ten-fold intersection case.

3.4 Special BBA

This is a special use of BBA because the optical axes are approximately parallel to the pavement and the bases move forward so a feature is imaged at different scales. It is a phototriangulation of terrestrial and frontal images as the photo scale varies significantly in a stereopair and along the sequence.

Together with some of the bases EOP, the base length BL (eq 4), and the base advancement BA (eq 5) were also constrained.

$$BL = \sqrt{(X'_0 - X''_0)^2 + (Y'_0 - Y''_0)^2 + (Z'_0 - Z''_0)^2} = 0.4m \quad (4)$$

$$BA = (0,5) [(Z_0^{k+} - Z_0^k) + (Z_0^{k+} - Z_0^k)] \approx 1,0m \quad (5)$$

where: one prime (') and two prime (") variables are, respectively, the left and the right 3D PC referred to LACRF; k and k+ refer to epoch t_k and t_{k+}.

Collinearity equations were the analytical math model, the indirect observation equation was the least squares estimation method, and the unit weight matrix synthesized the simplified stochastic model. Additionally we constrained the IOP calibrated elements, the first pair EOP of the sequence, and the BL and BA. The general solution is the equation (6):

$$\underline{x} = - (A^T P A + C^T P_c C)^{-1} \cdot (A^T P l + C^T P_c l_c) \quad (6)$$

where: A is de regular design matrix, P is the unit weight matrix related to the observations (photocoordinates), P_c is the weight matrix referred to the constraints, C is the "constrained design matrix", l and l_c are the vectors expressing the difference between computed and observed (l) and the pseudo-observed (l_c) values, respectively, along the iteration.

The *a posteriori* unit variance estimated the general behavior of all variables as a reference for the global quality of the least square adjustment. The very well-known equation (7) is

$$\sigma_0^2 = (\underline{v}^T P \underline{v} + \underline{v}_c^T P_c \underline{v}_c) / (n+n_c-u) \quad (7)$$

where: v and v_c are the residual and pseudo-residual vectors, respectively; n and n_c are the number of observations and pseudo-observations (constraints), and u the total number of unknowns (corrections to approximate parameters) to be estimated.

Considering a partial or the whole sequence, which is a partial or the full bundle block, accordingly, n is calculated by equation (8):

$$n = 2 \cdot tF - nTP \quad (8)$$

where tF is the total number of images (photos) in the block (partial or full), and nTP is the number of tie points.

The constraints, n_c, is calculated by equation (9):

$$n_c = (EOP+1) + tF \cdot IOP + nBL + nBA \quad (9)$$

where EOP is 6, IOP is equal 4 (f, x₀,y₀,k₁); nBL is the number of bases; and nBA the number of constraints related to base advances.

Finally, u is the total number of unknowns (eq 10):

$$u = (EOP + IOP) \cdot tF + 3 \cdot nTP \quad (10)$$

Concerning to the stochastic model, we can choose at least four distinct methods to assign values for the P (weight) matrix. According to the relative position of the tie point in the image plane, P can be assigned as a function of the distance to the center; or of the parallax of the conjugate tie points; or yet the identity matrix (P = I); or finally a diagonal matrix P = 1 / (pixel_size)² the one that we in fact used in the project.

3.4.1 Sequential EOP Estimation by Photogrammetric Bridging:

Every small block of a full sequence acts as a partial block that is solved by BBA of four to more bundles. This technique has been used to map streets and roads (Silva et al., 2000, 2003). We can start partial blocks of two bases; three bases (six images) considering two approaches: no base skipped such as we can have eight blocks (bases 1-2-3, 2-3-4 and so on until 8-9-10); one base skipped so we count four small blocks

(bases 1-2-3, 3-4-5, 5-6-7, and 7-8-9). Table 3 (Section 4, Experiments) summarizes more combinations.

3.4.2 BBA for the Whole Sequence: After the whole sequence has been covered by the partial blocks all EOP of each image were estimated through a sequential adjustment technique. In this project, it follows the BBA of the whole sequence in order to homogenize all EOP estimates. Here it was adapted to work in a semiautomatic manner.

4. EXPERIMENTS

4.1 Preparation

Two cameras Cannon T3i 600D were mounted on a rigid bar to form a stereo-pair with a base of 0.4m. An electronic device synchronized the shots. The original image size is 1920 x 1280 whose correspondent pixel size is 0.0115mm and 0.023mm for the 50% reduced image.

The calibration parameters of the camera system were estimated by the software CCv2011.05 (Galo, 2013) to estimate interior and relative parameters using the original image pairs taken in front of a 3D well determined coordinates target wall.

Ten bases were marked on the street pavement every one meter approximately to control de experiments. The stereo-camera was ported by a human operator on the marks. The sequential image pairs were taken along the sequence and saved as JPG format.

The LACRF is centered in the first pair left camera. A selection of 18 tie points in the first pair had their XYZ coordinates determined by topographic surveying using a total station. The first pair EOP was estimated by photogrammetric resection and they were constrained in the photogrammetric bridging process.

4.2 Tie point extraction and matching by SIFT

Considering the first two pairs (Figure 2), we continued to experiment with SIFT algorithm (initial tests are not reported in this article) to check the amount of the extracted key-points in the image plane (Table 1).

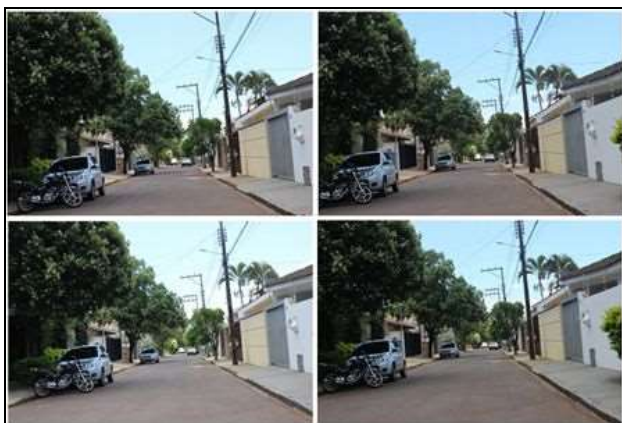


Figure 2 – The first two pairs: horizontal – pairs 1-2, 3-4; vertical: pairs 1-3, 2-4.

The amount of extracted key-points depends directly on the threshold value. Besides that the quantity extracted from forward pairs (1-3; 2-4) was larger than those from stereo-pairs (1-2; 3-4). Table 2 shows the results for threshold equal 0.6 and

50% reduced images (960 x 640 pixels), the same values used for all experiments. This confirms that SIFT is more efficient to match key-points from different scale forward images than similar scale lateral (stereoscopic) images, as declared by its creator.

| | | | | |
|--------|-------|-------|-------|---------|
| 0% red | 40% | 50% | 60% | 80% red |
| 13,520 | 5,172 | 3,509 | 2,364 | 685 |

Table 1 – Percentage of image reduction and the average of extracted key-points by SIFT

| | | | | |
|------------|-----|-----|------|------|
| Image pair | 1-2 | 3-4 | 1-3 | 2-4 |
| Matches | 681 | 847 | 1131 | 1028 |

Table 2 – Matched key-points from forward and lateral pairs.

Considering now all images of the sequence – 20 – the average number of extracted points was 3702 (the minimum 3289 in image 6 and the maximum 4513 in the image 12).

4.3 Arrangements of the experiments

Table 3 summarizes the experimented arrangements. Its first column identifies the experiment, B and I are the number of stereo-bases and images, accordingly, SB is the number of skipped bases, PB the amount of partial (sequential) blocks in the sequential BBA that is to say the photogrammetric bridging, and ATP the average amount of tie points used in PB.

| Exp | B | I | SB | PB | ATP |
|-----|----|----|----|----|-----|
| 1 | 2 | 4 | 0 | 9 | 269 |
| 2 | 3 | 6 | 0 | 8 | 140 |
| 3 | 3 | 6 | 1 | 4 | 138 |
| 4 | 4 | 8 | 0 | 7 | 80 |
| 5 | 4 | 8 | 1 | 4 | 80 |
| 6 | 4 | 8 | 2 | 3 | 82 |
| 7 | 5 | 10 | 1 | 3 | 47 |
| 8 | 5 | 10 | 1 | 2 | 49 |
| 9 | 10 | 20 | 5 | 2 | 15 |

Table 3 – Nine arrangements of the partial (sequential) blocks and the corresponding average amount of tie points.

Skipped bases are associated to the progress in the photogrammetric bridging. Exp 1 adjusted sequentially 9 partial blocks so that images 1 to 4 formed the first partial block, 3 to 6 the second, and so on until the last block was formed by images 17 to 20 with the largest average number of tie points. Exp 2 progressed with partial blocks formed by 6 images (1 to 6, then 3 to 8, 5-10, 7-12, and so on up to the last partial block formed by images 15 to 20). Exp 3 also used a partial block of six images but advancing such as 1 to 6, 5 to 10, 9 to 14, and 13 to 18. Exp 9 formed two partial blocks of five bases (10 images): 1 to 10 and 9 to 18 with only 15 average tie points.

| | | | |
|-----------|-----------|-----------|--|
| UL (3) | UC (2) | UR (3) | |
| CL (2) | CC (1) | CR (2) | |
| BL (3) | BC (2) | BR (3) | |

Figure 3 – Image plane sectors: U for upper, C for center, B for bottom, L for left, and R for right.

For better EOP estimation the distribution of tie points in the image plane must be balanced. We depicted the image plane in nine sectors (Figure 3) to count the amount of tie points in each sector (Table 4). A normalized index was computed based on the quantity of tie points and on weights (3, 2 or 1) according to the sector. The ideal index value would be 1.0 expressing that all tie points would be in the corner sectors. Low index expresses an unfavorable condition for estimating EOP by BBA.

| Sector | Weight | Exp 1 | Exp 5 | Exp 9 |
|--------|--------|-------|-------|-------|
| UL | 3 | 18 | 1 | - |
| UC | 2 | 16 | 6 | 3 |
| UR | 3 | 19 | 7 | - |
| CL | 2 | 37 | 14 | - |
| CC | 1 | 66 | 23 | 6 |
| CR | 2 | 24 | 4 | - |
| BL | 3 | 31 | 8 | - |
| BC | 2 | 36 | 12 | 3 |
| BR | 3 | 22 | 5 | 3 |
| Total | | 269 | 80 | 15 |
| Index | | 0.70 | 0.66 | 0.60 |

Table 4 – Average amount of tie points considering all images of the sequence and the average indices for the experiments.

We opted to present and discuss the results related to the extremes configurations (Exp 1 and Exp 9) and one in the middle (Exp 5).

5. BBA RESULTS AND ANALYSES

Table 5 presents all LACRF coordinates for the 20 PC estimated by BBA related to the experiments 1, 5 and 9. The only check condition is to compare the estimated Z_C coordinates to those approximately expected (multiple of one meter).

| Img. | Exp. 1 | | | Exp. 5 | | | Exp. 9 | | |
|------|--------|-------|-------|--------|-------|-------|--------|-------|-------|
| | Xc | Yc | Zc | Xc | Yc | Zc | Xc | Yc | Zc |
| 1 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2 | 0.40 | 0.00 | 0.00 | 0.40 | 0.00 | 0.00 | 0.40 | 0.00 | 0.01 |
| 3 | 0.00 | -0.04 | -1.03 | 0.00 | -0.04 | -1.05 | 0.00 | -0.04 | -1.06 |
| 4 | 0.40 | -0.04 | -1.02 | 0.40 | -0.04 | -1.05 | 0.40 | -0.03 | -1.04 |
| 5 | 0.02 | -0.11 | -1.99 | 0.04 | -0.07 | -2.05 | 0.04 | -0.07 | -1.99 |
| 6 | 0.42 | -0.12 | -1.98 | 0.44 | -0.07 | -2.00 | 0.43 | -0.07 | -1.95 |
| 7 | 0.00 | -0.18 | -3.03 | 0.01 | -0.14 | -3.06 | 0.01 | -0.11 | -2.96 |
| 8 | 0.40 | -0.19 | -3.03 | 0.41 | -0.13 | -3.03 | 0.42 | -0.10 | -2.91 |
| 9 | 0.01 | -0.31 | -4.04 | 0.01 | -0.23 | -4.04 | 0.06 | -0.17 | -3.96 |
| 10 | 0.41 | -0.31 | -4.04 | 0.41 | -0.24 | -4.07 | 0.48 | -0.18 | -3.97 |
| 11 | -0.01 | -0.35 | -5.00 | -0.02 | -0.27 | -4.98 | 0.05 | -0.20 | -4.92 |
| 12 | 0.39 | -0.35 | -5.02 | 0.38 | -0.27 | -5.04 | 0.46 | -0.20 | -4.96 |
| 13 | 0.06 | -0.39 | -6.03 | -0.01 | -0.33 | -6.04 | 0.10 | -0.26 | -6.04 |
| 14 | 0.46 | -0.39 | -6.03 | 0.39 | -0.33 | -6.04 | 0.50 | -0.25 | -6.01 |
| 15 | 0.06 | -0.45 | -7.01 | -0.01 | -0.39 | -7.02 | 0.18 | -0.30 | -7.04 |
| 16 | 0.46 | -0.46 | -7.01 | 0.39 | -0.39 | -7.02 | 0.58 | -0.29 | -7.02 |
| 17 | 0.08 | -0.50 | -7.97 | -0.02 | -0.46 | -7.98 | 0.21 | -0.34 | -7.98 |
| 18 | 0.48 | -0.50 | -7.96 | 0.38 | -0.46 | -8.00 | 0.62 | -0.34 | -7.98 |
| 19 | 0.11 | -0.60 | -9.00 | -0.03 | -0.53 | -9.00 | 0.26 | -0.40 | -8.97 |
| 20 | 0.51 | -0.61 | -9.00 | 0.37 | -0.54 | -9.03 | 0.66 | -0.40 | -8.97 |

Table 5 – Estimated terrain coordinates by the simultaneous BBA.

| | Exp 1 | Exp 5 | Exp 9 |
|-----------|--------|--------|--------|
| abs error | m | m | m |
| rms error | -0.009 | -0.025 | +0.014 |
| | 0.023 | 0.036 | 0.041 |

Table 6 – Absolute and rms errors for the Z_C .

The comparison let to the discrepancies and from them to the

absolute error (Figure 4) and to the root mean square (rms) error (Figure 5) for Z_C coordinates as shown in Table 6. The experiments had no control or check for the attitude parameters. The average standard deviation was computed for all oriented images (Table 7).

| | Exp 1 | Exp 5 | Exp 9 |
|----------|--------|--------|--------|
| ω | 0.0068 | 0.0072 | 0.0246 |
| ϕ | 0.0116 | 0.0086 | 0.0306 |
| κ | 0.0180 | 0.0220 | 0.0629 |

Table 7 – Estimated average standard deviation for the attitude parameters.

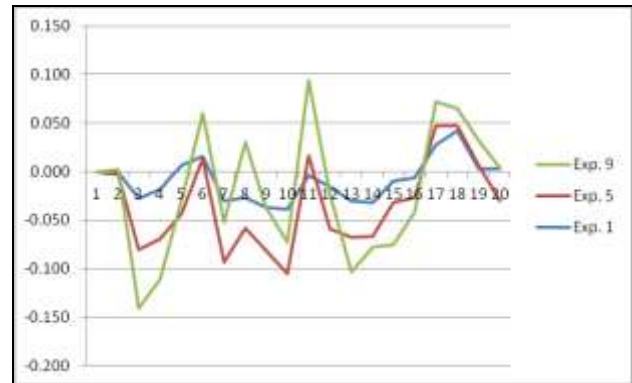


Figure 4 – Z_C absolute error at each PC.

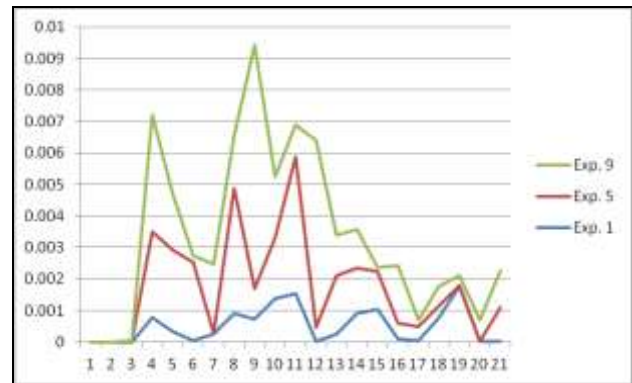


Figure 5 – Z_C rms error at each PC.

The analyses of the absolute and rms errors and also of the average standard deviation of the attitude parameters lead to the Exp 1 as the one that had the least error propagation as expected from the index 0.70 showed in table 4. Figure 6 and 7 are plots of the adjusted perspective centers corresponding the Exp 1.

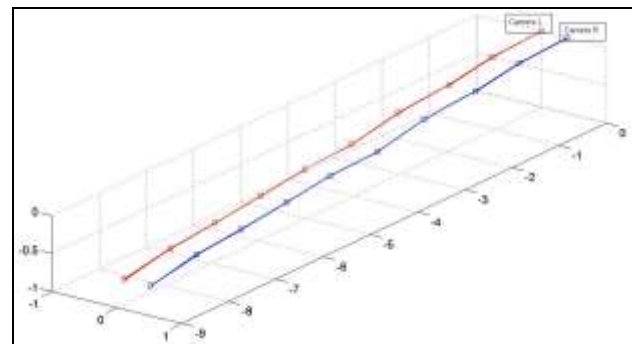


Figure 6 – Estimated positions of the PC plotted in the LACRF after BBA Exp 1.

The initial goal was to reach full automation to execute the entire project from the tie point extraction to the complete BBA. However, human intervention is still needed because decision issues related to outlier elimination, weight matrix balancing, integrating of sequential (bridging) to the complete sequence photo-triangulation to mention the most demanding at this point.

The coordinates of all PC had lower error propagation than for the attitude parameters. The approximately 1.0m constraint to the base advance may explain this.

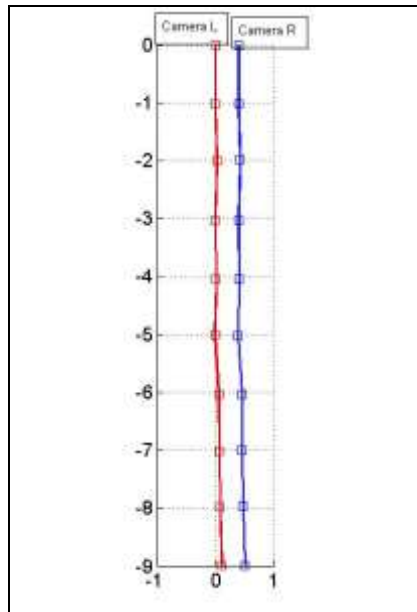


Figure 7 – Orthogonal plot of all PC estimated positions in the LACRF (Exp 1).

6. CONCLUSION

SIFT algorithm adapted well to this kind of stereo-pair sequence especially for the changing scale forward pairs. With the same algorithm we were able to automatize the tie point extraction and matching along the whole sequence. The amount of tie points obtained from the extracted key-points was sufficient for the BBA although in some images they are not uniformly distributed.

A fully automatic process is still on demand. In our particular application we need to automatize the photogrammetric process for the whole sequence when going from the bridging to the simultaneous and final adjustment.

Although a solution has been reached so that the hypothesis has not been denied, a complete confident result still depends on external data as we simulated an odometer measuring 1 m between the sequential bases. Therefore we also constrained the basis length besides the 1 m bases advancement. Both were necessary to avoid higher drift and torsion along the street image sequence.

7. FUTURE AND CONTINUED RESEARCH

Although SIFT algorithm has demonstrated efficiency to extract a large amount of key-points that are then corresponded to become tie points, the quality of such determination needs to be

evaluated in this particular application.

Besides the constraints related to base length and base advances proved to be effective to control the error propagation, it will be valuable if we can test and prove the contribution of low cost inertial sensors to help the attitude parameter estimation.

An algorithm to detect and localize outliers must be inserted in the photogrammetric chain.

The full photogrammetric process must be automatized so no human interaction is demanded. Decision rules based on statistical analyses may contribute to compose the weight matrix as function of parallax or Euclidean distance or other criteria.

New camera arrangements such as three cameras in line (collinear) forming a double stereo camera and three coplanar cameras are under construction to be tested. The coplanar arrangement means that the central camera in the collinear arrangement is lifted so the three PC makes a triangle such as all optical axes are parallel.

Another idea to be developed is based on the use of omnidirectional cameras so that a sequence of panoramic views may be investigated in the similar way we explored in this project.

ACKNOWLEDGEMENTS (OPTIONAL)

This work was supported by FAPESP (Proc. 03/00552-1), Unesp, and UTFPR.

REFERENCES

- Amami, M. M., Smith, M. J., Kokkas, N. 2014. Low cost vision based personal mobile mapping system. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Castelldefels, Spain, Vol. XL-3. 6p.
- Barbosa, R. L., Silva, J. F. C., Meneguete Jr, M., Gallis, R. B. A. 2007. Velocity estimation of a mobile mapping vehicle using filtered monocular optical flow. Proceedings. ISMMT2077. Padova. Italy.
- Brown, M., Lowe, D. G. 2002. Invariant features from interest point groups. In: British Machine Vision Conference. Cardiff, Wales. p.656-665.
- Galo, M. 2011. CC4_2013. http://www2.fct.unesp.br/docentes/carto/galo/_Softwares_Free/Calibracao/cc4_2013.exe (17 Oct 2014).
- Giachetti, A., Campani, M., Torre, V. 1998. The use of optical flow for road navigation. IEEE Transactions on robotics and automation, 14(1), pp. 34-48.
- Griessbach, D., Börner, A., Ernst, I., Zuev, S. 2010. Real-time dense stereo mapping for multi-sensor navigation. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII, Part 5, Newcastle upon Tyne, UK. pp.256-261.
- Habib, A. F. 1998. Motion parameter estimation by tracking stationary three-dimensional straight lines in image sequences. ISPRS Journal of Photogrammetry & Remote Sensing, 53,

pp.174-182.

techniques. *Journal of the Institute of Navigation*. Vol. 58. N. 2.
pp.127-139.

Lowe, D. G. 1999. Object recognition from local scale invariant features. In: *International Conference on Computer Vision*. Corfu, Greece. pp.1150-1157.

Lowe, D. G. 2004. Distinctive image features from scale-invariant keypoints. University of British Columbia. Computer Science Department. Vancouver. Canada. <http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf> (02 Nov 2012).

Moravec, H. 1981. Rover visual obstacle avoidance. In *International Joint Conference on Artificial Intelligence*, Vancouver, Canada, pp.785-790.

Randeniya, D. I. B., Gunaratne, M., Sarkar, S., Nasef, A. Calibration of inertial and vision systems as a prelude to multisensory fusion. *Science Direct. Part C*. p.255-74. 2008.

Roncella, R., Remondino, F., Forlani, G. 2005. Photogrammetric bridging of GPS outages in mobile mapping. *VIII Videometrics*, 5665, pp.308-319.

Rouzaud, D., Skaloud, J. 2011. Rigorous integration of inertial navigation with optical sensors by dynamic networks. *Navigation: J. of the Institute of Navigation*, Vol. 58, No. 2. pp.141-152.

Silva, A. R., Batista, J. C., Oliveira, R. A., Camargo, P. O., Silva, J. F. C. 1999. Surveying and mapping of urban streets by photogrammetric traverse. In: *International Workshop on Mobile Mapping Technology*. Bangkok. Proceedings. Vol. XXXII, pp.A5.1-5.

Silva, J. F. C., Barbosa, R. L., Junior, M. M., Gallis, R. B. A. 2007. Photogrammetric bridging using filtered monocular optical flow. *Proceedings. ISMMT2077*. Padova. Italy.

Silva, J. F. C., Camargo, P. O., Gallis, R. B. A. 2003. Development of a low-cost mobile mapping system: a South American experience. *Photogrammetric Record*, 18(101), pp.5-26.

Silva, J. F. C., Camargo, P. O., Oliveira, R. A., Gallis, R. B. A., Guardia, M. C., Reiss, M. L. L., Silva, R. A. C. 2000. A street map built by a mobile mapping system. In: *Proceedings of the XIX International Congress of ISPRS*, Amsterdam, Netherlands, pp.506-513.

Silva, J. F. C., Oliveira, R. A. 1998. Triangulation of a sequence of terrestrial digital images. In: *Proc. ISPRS Symposium on data integration: systems and techniques*. Cambridge, UK, pp. 273-277.

Silva, J. F. C., Barbosa, R. L., Júnior, M. M., Gallis, R. B. A. 2007. Photogrammetric bridging using filtered monocular optical flow. *Proceedings. ISMMT2077*. Padova. Italy.

Tao, C., Chapman, M. A., Chaplin, B. A. 2001. Automated processing of mobile mapping image sequences. *Journal of Photogrammetry & Remote Sensing*, 55, pp. 330–346.

Tao, C., Li, J. 2007. *Advances in Mobile Mapping Technology*. Taylor & Francis, London (ISPRS Book Series).

Veth, M. J. 2011. Navigation using images: a survey of