AN APPROACH FOR AUTOMATIC STEREO MODEL GENERATION USING NON-METRIC DIGITAL IMAGES

Roberto da Silva Ruy
Antonio Maria Garcia Tommaselli
Elaine Reis Costa
Laysson Guillen Albuquerque
Mauricio Galo

Dept. of Cartography, São Paulo State University at Presidente Prudente
(rruy, elaine, laysson)@pos.prudente.unesp.br
(tomaseli, galo)@prudente.unesp.br

KEY WORDS: Photogrammetry, Acquisition, Correlation, Orientation, Reconstruction, Automation.

ABSTRACT

The evolution of non-metric digital cameras and its integration with direct orientation sensors (GPS/INS), make feasible some applications that require fast mapping, like thematic and cadastral mapping, disasters management and environment monitoring. However, the accuracies of the GPS/INS sensors are usually not enough to generate a stereo-model without vertical parallax, hindering the stereoscopic visualization of the scene and affecting the 3D reconstruction. This paper presents an automatic methodology for removing the vertical parallax in the model, based on a modified coplanarity model. Some tie points are automatically measured in the model using area-based correspondence methods, integrating methods to reduce the search space and an *a priori* analysis of the matching areas to increase the robustness of the process. After the EO parameters adjustment, the images are normalized through an epipolar resampling, in order to provide a suitable stereoscopic visualization and to facilitate the process of automatic Digital Terrain Model (DTM) generation. A methodology for DTM generation, based on the adjusted EO parameters and epipolar images is presented as well. Some experiments in a test area were performed, and the results obtained are discussed, showing the effectiveness of the proposed approach.

1. Introduction

With the evolution of digital image acquisition devices, combined to the development of sensors for direct positioning and attitude determination, some applications that need fast mapping could benefit, like thematic and cadastral mapping, disasters management and environment monitoring. These systems are based on digital cameras and sensors for direct georeferencing (GPS+INS) and are used in the fast acquisition of georeferenced data.

A key problem with digital cameras is the sensor resolution and the sensitive area of the sensor, when compared to convencional metric cameras in Photogrammetry, making necessary many frames to cover a reference area on terrain. Thus, the complexity of exterior orientation (EO) determination by aerotriangulation increases, because many control and tie points have to be measured in the process, resulting in expensive ground control surveying.

One of the alternatives to solve this problem is the integration of GPS and INS for direct determination of the EO parameters. However, the parameters supplied by these sensors are under the required accuracy to generate a stereo model without vertical parallax [1, 2, 8], making difficult the stereoscopic visualization of the scene and affecting 3D reconstruction and DTM generation.

In this paper, an automatic approach to remove the residual vertical parallax in the model is presented, based on the reprocessing of the orientation parameters, supplied by the integrated GPS/INS sensor. The coplanarity model was modified and the twelve EO parameters are adjusted by the Least Square Method (LSM) with constraints. After that, the images are normalized by epipolar resampling, providing a suitable stereoscopic visualization of the model and improving the correspondence process for automatic Digital Terrain Model (DTM) generation. The methodology for DTM generation, based on the adjusted EO parameters and epipolar images, is presented as well.

2. Direct Georeferencing

Direct georeferencing consists on the determination of the camera exterior orientation parameters during exposure, using the integrated GPS and INS sensors. The Global Positioning System (GPS) is used for camera perspective center location. In the inertial navigation system (INS), an Inertial Measured Unit (IMU), composed by accelerometers and gyroscopes, determines the camera attitude during the flight.

The system integrating GPS and INS sensors is a powerful tool since the inertial system can help GPS by providing accurate initial position and velocity information in case of signal loss. Even during satellite outages, when the number of visible satellites drops below four, INS will provide continuous trajectory information. On the other hand, the high absolute performance obtained by GPS can help the inertial navigation system with accurate estimates on the current behaviour of its error statistics. If Kalman filtering is used in traditional navigation approaches, the internal INS errors are modelled as gyroscope drifts and accelerometer offsets, using GPS information [3].

One point that must be considered in direct georeferencing is the reference system. The orientation angles from GPS/INS are not comparable to the photogrammetric angles (ω, φ, κ). Because the integrated GPS and INS solution was originally designed for navigation purposes, the computed attitudes are interpreted as navigation angles (roll – r, pich – p, yaw - y). These angles are obtained in a local navigation system, whose origin is located in the center of the inertial sensors axes triad. Since the INS is moving relatively to the earth's surface, this local frame is not constant but moves with time. Consequently, it is necessary to transform the navigation angles to photogrammetric attitudes, via the Cartesian earth-centered earth-fixed coordinate system [3].

In image orientation there is an important problem: the overall system calibration. Any discrepancies between the assumed mathematical model used and the true physical reality during image exposure will cause erros in object point determination. In direct georeferencing the interior geometry of the images is of crucial important, because GPS/INS provides direct measurements of the true physical camera position and orientation during exposure, whereas in the bundle adjustment the values of the exterior parameters are estimated indirectly. A great advantage of direct georeferencing is to minimize or even eliminate the need for control points in exterior orientation process, mainly in areas with difficult access.

Although GPS and INS systems have high cost, the direct sensor orientation during the flight can be advantageous, due to:

- Fast determination of the EO parameters, because the aerotriangulation can be eliminated or totally automatized;
- It allows to observe time-variant phenomena, that requires fast cartography (flood, sign of petroleum, etc);
- There is no restriction to flight directions, and corridors can be easily covered without making a block of images;
- It allows the automatic control of the camera exposures;

The accuracy of the EO parameters provided by the integrated GPS and INS depends on the system. Generally, systems used for mapping present accuracies around 0.1-0.3m for camera perspective center location (X_0, Y_0, Z_0) and 1-3' for attitude angles (ω, φ, κ).

3. Stereo model generation

The methodology proposed in this paper for automatic stereo model generation is based on the reprocessing (adjustment) of the exterior orientation parameters provided by the integrated GPS and INS, removing the vertical parallax in the stereo model. The input data are the aerial digital images, with its camera calibration parameters, and the EO parameters supplied by GPS/INS, with their accuracies.

In general, three steps are involved in this methodology: automatic tie point measurement; reprocessing of the EO parameters and epipolar resampling.

3.1 Automatic Tie Points Measurement

The adjustment of the EO parameters using Least-Squares Method (LSM) requires some tie points that can be automatically measured in the model. In this approach one single tie point is enough to refine these parameters, since all the twelve EO parameters are considerated as weighted constraints. Nevertheless, the adjustment process will be more reliable if more tie points are included in this process, since points measured erroneously can be removed by the residual analysis.

An approach for automatic tie point measurement was implemented, using an area-based correspondence method, employing a correlation function and LSM matching (for details, see [10]). To reduce false correspondences the search areas are marked as polygons, from which scan lines are generated; moreover, an *a priori* analysis of each reference window, based on the Förstner interest operator, is performed, as follows.

Scan-lines are used to reduce the search space for the area-based correspondence methods (Figure 1). In this process the image point is projected to the terrain by inverse collinearity equations, using the EO parameters supplied by GPS/INS and approximate elevation of the point (P) supplied from any available DEM, e.g., SRTM. Since the value of elevation is approximated, an uncertainty (Δz) is associated to this elevation. The maximum (U) and minimum (L) values of this range are projected to the right image by collinearity equations, generating an epipolar line (u" – l"). Instead of using a squared search area, a polygon around the epipolar line is generated by determining the uncertainty of the endpoints using covariance propagation, considering the covariances of the EO parameters. Figure 1 depicts this concept.

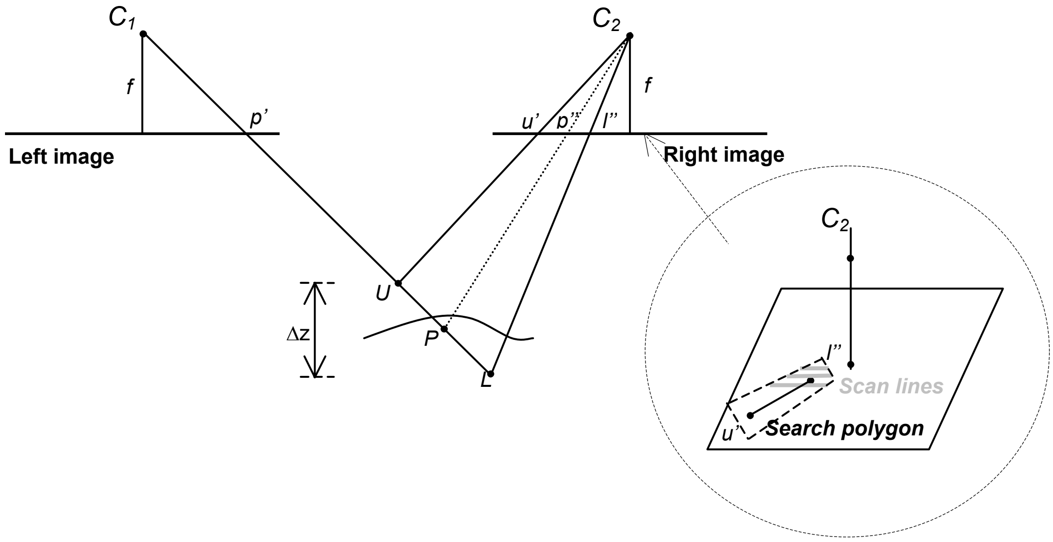


Figure 1: Scan lines to reduce the search space in matching process.

The correspondence process is preceded by an analysis of the matching areas, to avoid unsuitable (homogeneous) areas. This *a priori* analysis is based on the computation of the covariance matrix (∑) of the radiometric translation components of the matching areas in the reference template. This matrix is a function of the noise variance and the gray level gradients on the template. Thus, the area-based matching (cross-correlation technique and LSM) is applied only for those areas with acceptable translation covariance matrix. The following three steps are derived from [5], where the covariance matrix of the estimated shifts is the covariance matrix of the Förstner interest operator [4]:

- *Computation of the noise variance (σ_n^2) based on the template variance (σ_r^2) and the maximum correlation coefficient (ρ_{12}):*

$$\sigma_n^2 = \frac{\sigma_r^2(1-\rho_{12})}{\rho_{12}} \tag{1}$$

The maximum correlation coefficient (ρ_{12}) is a matching value theoretically correct for the images. Since this approach is totally automatic, this coefficient was defined as 0.8, based on empirical tests.

- *Computation of the normal equations coefficient matrix (N) based on the gray level gradients of the template:*

Considering only radiometric translations between the images, the linearized model for the correspondence process is given by equation 2, where the translation parameters (Δx, Δy) can be estimated in an adjustment by the Least Squares Method:

$$\Delta g = -f_r.\Delta x - f_c.\Delta y + n \tag{2}$$

where $n =$ noise;
$f_r \; e \; f_c$ = row (r) and column (c) gradients.

The normal equation coefficient matrix (N) is given by:

$$N = \begin{pmatrix} \sum_{i=1}^{m} f_{r_i}^2 & \sum_{i=1}^{m} f_{r_i} f_{c_i} \\ \sum_{i=1}^{m} f_{r_i} f_{c_i} & \sum_{i=1}^{m} f_{c_i}^2 \end{pmatrix} \quad (3)$$

- *Computation of the accuracy of the translation parameters in function of the noise variance (σ_n^2) and the normal equations coefficient matrix (N):*

The accuracy of the translation parameters (Δx, Δy) is given by:

$$\Sigma_{\Delta x,\Delta y} = \hat{\sigma}_n^2 \cdot N^{-1} = \hat{\sigma}_n^2 \cdot \frac{1}{n_{11}n_{22} - n_{12}^2} \begin{pmatrix} n_{22} & -n_{12} \\ -n_{21} & n_{11} \end{pmatrix} \quad (4)$$

where $\hat{\sigma}_n^2$ = the estimated noise variance considerated equivalent to the computed noise variance (σ_n^2);
$n_{11},\ldots,n_{22}$ = elements of the matrix *N*.

Then:

$$\hat{\sigma}_{\Delta x} = \hat{\sigma}_n \cdot \sqrt{\frac{n_{22}}{n_{11}n_{22} - n_{12}^2}} \;,\; \hat{\sigma}_{\Delta y} = \hat{\sigma}_n \cdot \sqrt{\frac{n_{11}}{n_{11}n_{22} - n_{12}^2}} \quad (5)$$

Based on the previous equations it can be seen that the covariance matrix of the radiometric translation components and, consequently, the standard deviation of the parameters, depends on: the noise variance; the number of pixels of the template, since the standard deviation decreases linearly when the template size increases and; the gray level gradients in the template, which indicate the presence of edges.

Considering the noise variance as a constant in the overall image, it is possible to determine the locations that most probably will provide high accuracy in the correspondence process, based only on the template data.

In this work the thresholds for unsuitable areas were empirically established after some tests with real data. These tests were performed in various conditions, from homogeneous areas to areas with high variance. Thus, based on these experiments, the followings thresholds were considered:

Template variance (σ_r^2) < 300 or $\sum$ trace > 0.05

If the template is considered doubtful, another template (translated by five horizontal pixels) is selected. This approach is repeated until the template is considered suitable.

3.2 Reprocessing of the EO parameters

The methodology proposed in this paper is based on the reprocessing of the exterior orientation (EO) parameters provided by the GPS/INS sensors, using a modified coplanarity model (eq. 6). In this model, the coplanarity equation is rewritten considering all the twelve EO parameters, which are treated as weighted constrains. These constraints come from the estimated covariance matrix of the data provided by the sensors of position (GPS) and attitude (INS).

$$\left(X_{L_2} - X_{L_1}\right) * \left(v_1 \cdot w_2 - v_2 \cdot w_1\right) + \left(Y_{L_2} - Y_{L_1}\right) * \left(u_2 \cdot w_1 - u_1 \cdot w_2\right) + \left(Z_{L_2} - Z_{L_1}\right) * \left(u_1 \cdot v_2 - u_2 \cdot v_1\right) = 0 \quad (6)$$

with:

$$\begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} = M_1^T \begin{bmatrix} x \\ y \\ -f \end{bmatrix}_1 \quad \text{and} \quad \begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = M_2^T \begin{bmatrix} x \\ y \\ -f \end{bmatrix}_2 \quad (7)$$

where:
$\left(X_{L_1}, Y_{L_1}, Z_{L_1}\right)$: camera perspective center location for the left image in object space coordinate system;

$(X_{L_2}, Y_{L_2}, Z_{L_2})$: camera perspective center location for the right image in object space coordinate system;

M_1 : rotation matrix for the left image;

M_2 : rotation matrix for the right image;

$(x, y, -f)_1$: image coordinates of a point in the left image;

$(x, y, -f)_2$: image coordinates of a point in the right image.

For reprocessing of the EO parameters, an adjustment by LSM with nonlinear conditions (See [6]) was implemented, since observations and parameters are combined in a non-explicit equation.

$$F(L_a, X_a) = 0 \tag{8}$$

Since the coplanarity model is represented by one equation for each measured point in the images, then:
- Number of equations (r) = number of measured points in the images (n_p);
- Number of observations (n) = 4 * number of measured points in the images (n_p);
- Number of constraints (s) = number of parameters (u) = 12.

With $\boldsymbol{S = r + s - u}$ degrees of freedom it is necessary only one equation for the estimation with the Least Squares Method, since all the exterior orientation parameters are considered as weighted constraints. For details of this approach, see [9].

3.3 Epipolar Resampling

This process is carried out for removing the vertical parallax and to provide a suitable stereoscopic visualization of the model, with the resulting improvement also of the correspondence process, because the conjugate entities are resampled to correspondent rows in each image (See Figure 6).

Transforming the original images in their normalized positions, requires two steps [7]:
- The images are transformed for their vertical positions, by computing the rotation matrixes;
- Starting from vertical images, the images are transformed for their normalized positions, by computing the base rotation matrix.

4. DTM generation

For digital terrain model generation, a grid of conjugated points (Figure 2) are automatically measured in the model, using the approach described in section 3.1. Based on the epipolar resampled images, the coordinates of the correspondent points and the adjusted EO parameters, the 3D coordinates of the grid points are computed by photogrammetric intersection (See [7]).

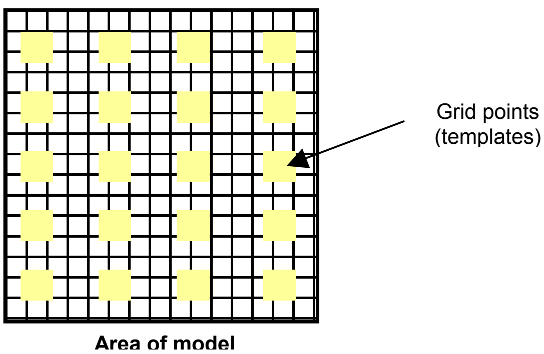


Figure 2: Grid of conjugated points for DTM generation.

The steps for computing the 3D coordinates of the grid points measured in the stereomodel are summarized in Figure 3.

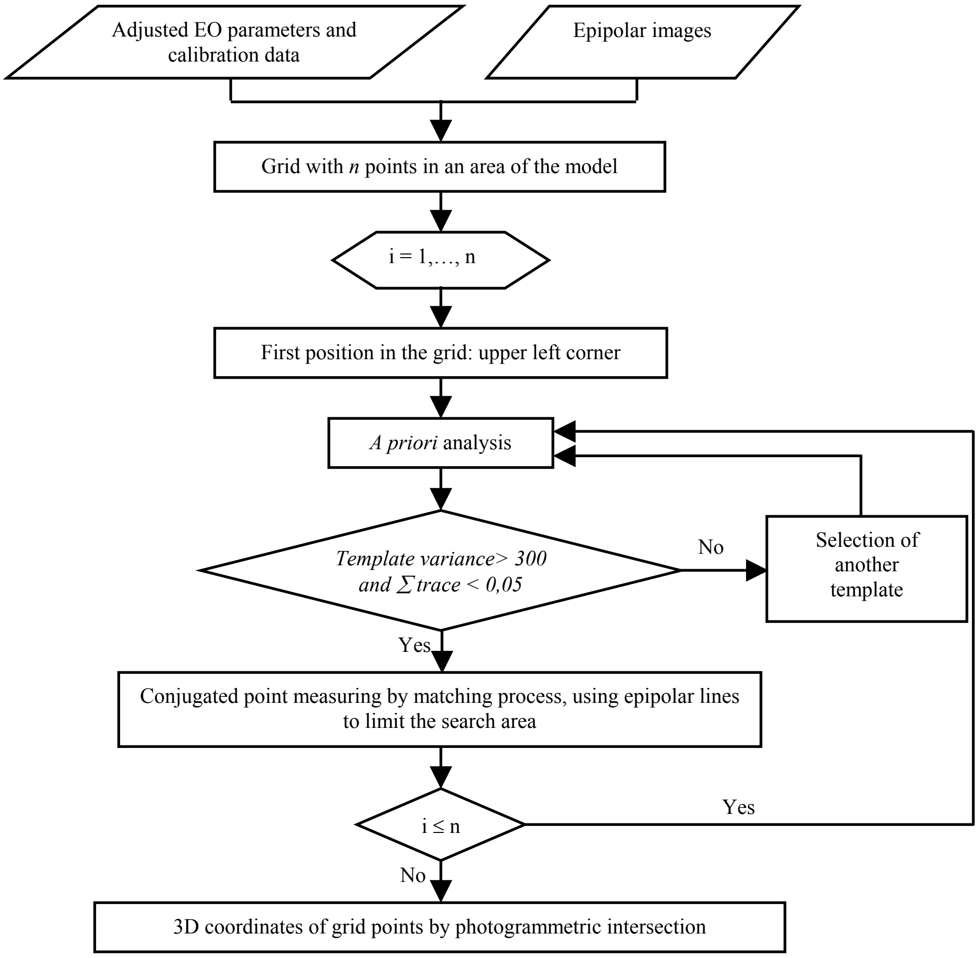


Figure 3: Summary of the 3D reconstruction for DTM generation.

5. Experiments

5.1 Stereo Model Generation

Based on the methodology presented in section 3, some computational programs in C/C++ languages were developed. Some experiments were performed to assess the vertical parallax effect and the accuracy of check points coordinates on terrain, before and after reprocessing the EO parameters.

In this experiment a pair of images taken by the KODAK PRO 14n digital camera was used, over the Unesp campus in Presidente Prudente, SP, Brazil. The images were obtained from a flight height of 1550m, with a scale of 1:30000, image resolution of 4500 x 3000 pixels and pixel size of 8µm. The nominal camera focal length was 50mm.

Since this flight was performed without direct orientation sensors, the EO parameters were estimated in an aerotriangulation process, using a digital photogrammetric workstation (Socet Set – BAE). In order to simulate direct orientation using GPS and INS, some random perturbations were introduced in the position and attitude parameters. Standard deviations of 0.5m and 10’ were considered for the perspective center location and the attitude angles, respectively. These values were used considering low cost direct orientation sensors.

The observations (tie points coordinates) were automatically measured in the model, using the approach presented in section 3.1. For these experiments, nine points were automatically measured, but two of these points were eliminated after the residual analysis (residuals higher than 0.5 pixel).

A set of experiments with this level of accuracy was performed and the results are summarized in Table 1. The Root Mean Square (RMS) values for five check points, and the vertical parallax (py) in the model are presented considering three conditions: A) after the aerotriangulation; B) with the EO parameters degraded (before reprocessing) and; C) after reprocessing of the EO parameters.

	CONDITIONS		
	A	**B**	**C**
RMS_X (m)	0.68	1.32	1.60
RMS_Y (m)	0.63	1.55	3.30
RMS_Z (m)	1.67	9.47	4.27
py (mm)	*0.0223*	*0.0834*	*0.0026*

Table 1: RMS values of check points and vertical parallax (*py*) in the model.

It was verified that the vertical parallax was reduced from 0.0834mm (10.43 pixels) to a mean value of 0.0026mm (0.33 pixel), which is smaller than the parallax after the aerotriangulation (2.78 pixels).

The XY components showed higher errors after the reprocessing of the EO parameters (See Table 1), probably due to the lack of accurate interior orientation parameters. Besides that, since the coplanarity equation is independent of object space data, the images are adjusted for providing a better solution concerning the stereo visualization of the model, like a relative orientation.

The RMS error of the computed elevations was reduced, when compared to situation B, but it still can be considered a high value, due the lack of knowledge of the focal length.

In these tests, systematic discrepancies on the check-points coordinates were detected. The resulting model was translated and rotated in the space, with respect to the reference system defined by control points. Then, if some control points (minimum of three) are available, a seven-parameter transformation can be applied to correct the ground coordinates, although this is out of the objectives of the direct georeferencing methodology.

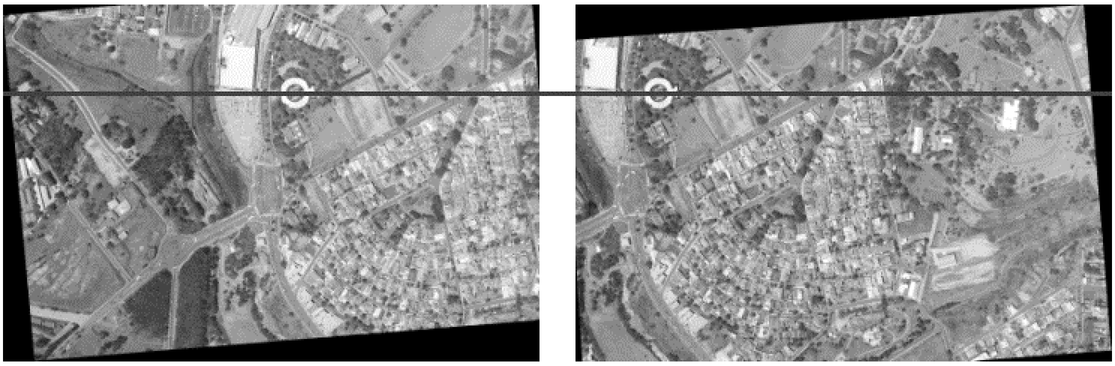
The visual effect of the vertical parallax in the model can be seen in Figure 4, which shows a part of the model before and after the reprocessing of the EO parameters, with the normalized images.

(a) **(b)**

Figure 4: Vertical parallax in the model: (a) before and (b) after the reprocessing of the EO parameters and epipolar resampling.

The influence of the vertical parallax on stereoscopic visualization of the model can be seen in Figure 4. In Figure 4a stereoscopic visualization is difficult because of the high vertical parallax value (about ten pixels). After the reprocessing of the EO parameters and epipolar resampling, however, the vertical parallax was significantly reduced (0.33 pixel), providing a suitable stereoscopic visualization (Figure 4b). Furthermore, after the epipolar resampling, the conjugated entities in the normalized images are in the same row, as showed in Figure 5.

(a)

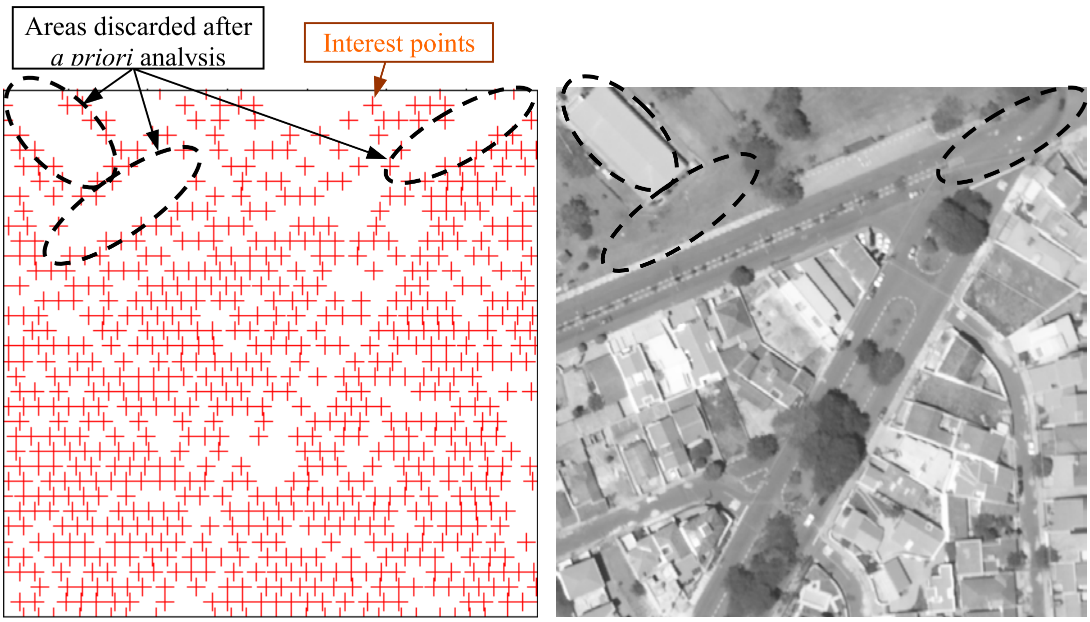


(b)

Figure 5: (a) original images (b) normalized images: the conjugated entities are resampled to the same row.

5.1 DTM Generation

The DTM of a select area in the model can be generated from the normalized images, presented in Figure 5. A grid of interest points is automatically measured in the model using the matching process described in section 3.1. Points presenting high variances are eliminated after the *a priori* analysis of the templates. In Figure 6, examples of areas with homogeneous gray level, that were discarded in the matching process, are showed.



(a) (b)

Figure 6: Example of points used for DTM generation: (a) Interest points with discarded areas; (b) Resampled gray level image.

For these points accepted in the matching process, the coordinates of the object space are obtained, and the 3D object space coordinates are computed by photogrammetric intersection [7]. The process for DTM generation is still under development and several filtering strategies are being implemented. Two approaches are being studied: a quality control that considers both the correlation coefficient and the covariance matrix of the translation; and the previous labelling and segmentation of the image, this requiring RGB or IR images.

6. Conclusions

This paper presented an approach for 3D model reconstruction, by removing the vertical parallax in the model, when considering EO parameters from integrated GPS and INS.

In all studied cases the proposed approach was efficient in removing the existing vertical parallax, due to the level of accuracy of the directly measured EO parameters.

The experiments with a KODAK PRO 14n camera were limited because of the lack of IO data, but even when using the nominal focal length (50mm) good results were achieved. Some tie points were eliminated, probably due to the effect of non-compensated lens distortion.

It is important to mention that this approach produces a “near leveled model”, because the coplanarity model is independent of control points. It is rather an attempt to provide a better solution concerning the stereo visualization. If low cost sensors were used for direct georeferencing it will be necessary a geometric transformation to the object space to reduce the discrepancies, according to the experiments performed with the PRO 14n camera.

7. Acknowledgements

The authors wish to thank to Fapesp (Fundação de Amparo à Pesquisa do Estado de São Paulo), to CAPES (Coordenação de Aperfeiçoamento do Pessoal do Ensino Superior) and to CNPq (Conselho Nacional de Desenvolvimento Científico) that provided grants and scholarships for the development of this project.

8. References

[1] Brzezinska, D. A. G., 1999. Direct Exterior Orientation Of Airborne Imagery with GPS/INS System: Performance Analysis. *Navigation*, Vol. 46, No. 4, pp. 261-270.

[2] Colomina, I., 1999. GPS, INS and Aerial Triangulation: What is the Best Way for Operation Determination of Photogrammetric Image Orientation. *IAPRS*. München.v. 32, Part. 3 – 2W5. September 8-10. pp.121-130.

[3] Cramer, M.; Stallmann, D., 2002. On the use of GPS/inertial Exterior Orientation Parameters in Airborne Photogrammetry. *OEEPE – Integrated Sensor Orientation*: Test Report and Workshop Proceedings. Frankfurt, Kartographie und Geodasie.

[4] Förstner, W., 1986. A feature based correspondence algorithm for image matching. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXVI (3/3), pp. 150-166.

[5] Haralick, R; Shapiro, L.G., 1993. *Computer and robot vision*. Addison-Wesley Publishing Company. v.2.

[6] Mikhail, E. M.; Ackerman, F. *Observations and Least Squares*. New York: IEP, 1976. 497p.

[7] Mikhail, E. M.; Bethel, J. S. Mcglone, J. C., 2001. *Introduction to Modern Photogrammetry.* Inc. New York : John Wiley & Sons, 479p.

[8] Mostafa, M. M. R.; Schwarz, K. P., 2001. Digital Image Georeferencing from a Multiple Camera System by GPS/INS. *ISPRS Journal of Photogrammetry & Remote Sensing,* 56, pp. 1-12.

[9] Ruy, R. S., 2004. *Geração Automática de Estereomodelos a partir de Imagens Digitais Georreferenciadas*. Dissertação (Mestrado em Ciências Cartográficas) – Universidade Estadual Paulista, Presidente Prudente.

[10] Schenk, T., 1999. *Digital Photogrammetry.* The Ohio State University: TerraScience, v. 1.