Orbital image correction for multiple models of geometric transformations

Emiliano Ferreira Castejon [1]
Carlos Henrique Quartucci Forster [2]
Leila Maria Garcia Fonseca [1]
Thales Sehn Korting [1]

[1] Instituto Nacional de Pesquisas Espaciais - INPE
12227-010 - São José dos Campos - SP, Brasil
{castejon,leila,tkorting}@dpi.inpe.br

[2] Instituto Tecnológico da Aeronáutica - ITA
12228-900 - São Jose dos Campos, SP – Brasil
forster@ita.br

Abstract. This paper describes a method for geometric correction of orbital images supporting multiple models of geometric transformations. It is based on a previous method and represents a step forward by incorporating a new form of combining spectral information with the information from the geometric deformation between a pair of images. From this point it is possible to define homologous points used for the correction of geometry and positioning.

Keywords. Image processing, Geometric correction, Registration, Remote Sensing.

1. Introduction

It is known that the problem of automatic satellite image registration is a difficult task because of their characteristics. Many approaches for image registration have been provided in the literature but none of them has been capable of effectively solve the problem as shown by the survey Zitová and Flusser, (2003). Since the performance of a methodology is dependent on specific application, sensor characteristics, and the nature and composition of the imaged area, it is unlikely that a single registration scheme will work satisfactorily for all different applications Fonseca and Manjunath, (1996).

This paper proposes a method for geometric correction based on a previous work from Fedorov, (2002). This method starts from two images: a *reference image*, considered to be geometrically correct; and the one to be corrected, known as the *adjust image*. The method searches for radiometrically similar positions, in both images and use them to build a geometric transformation that models the deformation. Finally this model is used to correct the entire adjust image. This work goes a step forward by incorporating a new geometric filter approach used in conjunction with a more stable method to detect interest points.

The organization of this paper presents the method in a way that allows the reader to reproduce and verify correctness and effectiveness. Section 2 and 3 describe how to localize the features and points of interest. Section 4 presents the definition of geometric transformation parameters. In section 5 we provide some results of our method, followed by conclusions in section 6.

2. Localization of features of interest

The objective of the initial step from this method is the localization of features of interest at both images, i.e. those with large size and which would probably be at the two images. These features will be used for the definition of homologue points between them. The types of features most commonly used are points, polygons and lines. Each type is best applied to a certain sort of image data, as shown on Castejon et al, (2009). Since the data used are from the same origin, punctual features were used, characterized by corners and borders of image objects. These types of features could give good results at the next step (localization of homologous points), because they can produce good local similarity coefficients, even under

the effect produced by local deformation, inherent to the model of geometric transformation which maps both images.

The followed technique is based on a variation of what was originally proposed in Moravec, (1997). Its effectiveness and efficiency when applied to remote sensing data was confirmed at USGS (2010), for the registration of Landsat GeoCover 2000 data. The Moravec operator acts similarly to a linear filter calculating the directional variance of each point by analyzing its neighborhood under a certain radius *m* centered in the point in question (defined here as Moravec's window).

The result of the application from this operator at each of both images (images I1 and I2) are two surfaces with directional variance. At these surfaces those points with highest values are points of interest. Computationally it is not viable to use all points in the process, and due to that, for its optimization, a method is needed to select which are the best points among all those localized. A variation of the method described in USGS (2010) was applied. This method considers the distribution homogeneity over the entire image. This is an important factor since the concentration of points of interest in only specific regions from the image tends to generate parameters of a geometric transformation model which does not correctly covers the other image sections where there is not a great concentration of points, even considering that these points have good geometric quality and good directional variance values.

3. Localization of homologous points

The product of earlier steps is a set of points indicating positions in interest features at each one of the two images and the respective values of directional variance. For a possible association between a point at image A and B, it is necessary to create a local descriptor capable of representing that point. The comparison is then made over the descriptor which could be composed by a set of contextual characteristics from the point. An example would be: the relation and positioning of the point related to its neighborhood.

In the literature there are several forms to define a local descriptor of a point. One example is the work of Lowe, (1999), where the concept of directional histograms is defined, which condensates the information of orientation and magnitude of points that compose it. The descriptor is then defined by the concatenation of these histograms. According to Castejon et al, (2009), a good cost/benefit ratio was verified if the local descriptor of a point would be assembled by the use of the own values of neighborhood points under a predefined radius. The radius parameter determines the size of the descriptor and it has a direct influence on the result. Fonseca, (1999) stated that when using images of sensor CCD/CBERS, 21 (pixel units) is an excellent value for the diameter. To achieve rotational invariance, a rotation of the image data covered by the interest point neighborhood is performed before the related descriptor generation like showed in Fonseca and Kenney, (1999).

Given two descriptors representing two interest points it is possible to use a method of comparison to find their degree of similarity. For those descriptors assembled with the use of values from neighborhood points, a commonly used measure is the value of common information, defined as a function of the entropy of each image and of the joint entropy between two images, as shown in Viola and Wells, (1997). Based on the results from Castejon et al, (2009), the chosen strategy was to use the correlation coefficient among two descriptors, as shown in Pratt, (1974).

To use only the correlation coefficient value would not always bring good results and so o reduce the occurrence of incorrect matches due to bad correlation, the directional variance values, created during the phase of maxima points localization, are combined with the correlation coefficient value, generated from the comparison of descriptors. The created information, defined now as the normalized weight of the control point, is used in the next

step to help the removal of wrong points (outliers). The calculus of the weight is made as shown in equation 3.5 and the normalized weight calculus is presented at equation 3.6.

$$P = \left\{\left(\frac{V_1 - MinV_1}{MaxV_1 - MinV1_1}\right) + \left(\frac{V_2 - MinV_2}{MaxV_2 - MinV_2}\right)\right\} \times \left(\frac{C - MaxC}{MaxC - MinC}\right) \quad (3.5)$$

$$P_{norm} = \left(\frac{P - MinP}{MaxP - MinP}\right) \quad (3.6)$$

Where, P is the non-normalized weight associated to the control point, V_1 is the value of directional variance for the maxima point at image A, $MinV_1$ is the minimum value of directional variance, considering all maxima points found for image A, $MaxV_1$ is the maximum value of directional variance, considering all maxima points values found for image A. V_2 is the value of directional variance for the maxima point at image B, $MinV_2$ is the minimum value of directional variance considering all maxima points found at image B, $MaxV_2$ is the maximum value of directional variance considering all points of maxima found at image B. C is the correlation coefficient associated to the point of control, $MaxC$ is the maximum value of the correlation coefficient considering all control points, $MinC$ is the minimum value of the correlation coefficient considering all control points, and P_{norm} is the normalized weight associated to a control point.

4. Removal of outliers and definition of transformation parameters

The result of the previous step is a set of control points in indexed coordinates (column, line) of homologous points at images A and B, as well as a set of normalized weights associated to each one of them. As mentioned previously in this study, matches that would generate incorrect control points are possible. The presence of these incorrect points can cause some level of interference when the parameters of the geometric transformation model are defined; and it can cause deformations on the final product generated.

There are several studies on the problem of outlier removal. One of the most used, due to the low computational cost, is the RANSAC method described at Fischler and Bolles, (1987). It is a method which performs random sub-optimal selections on the initial set of control points and iteratively refines the parameters of a geometric transformation, minimizing an error parameter for the chosen points. This method was applied at Fedorov, (2002) to determine the parameters of an affine transformation, given an initial set of control points at each one of the two images used. One of the requirements of this study is the capacity to work with different models of geometric transformations, i.e. with a variable amount of parameters. Another requirement is the use of weights for the control points, in order to guarantee that the best points are those which define the final geometric transformation. Considering the control points generated at the previous step are in its majority correct, since they passed through an exhaustive procedure of comparison, we chose an iterative method which works inversely to the RANSAC method.

This method estimates the parameters of initial geometric transformation parameters using all the points and then performs the refinement by the iterative removal of the worst control points. Eventually discarded points could be reutilized, because it is possible that at certain intermediate iterations, points which were first considered poor, comply with the criterion of predefined error. The reutilization of points minimizes the tendency to have a point agglomeration improving its distribution. One considers that, given a generic model of geometric transformation T and a minimum number of required control points, it is possible to calculate its parameters as well as the parameters of the respective inverse transformation T-1. The same is true for larger amounts of control points (for these cases an estimation method

must be used, such as the method of least squares described in Umeyama, (1991)). One considers that a geometric transformation T and the respective T-1 don't have outliers, if they satisfy two criteria: The average square error of the two transformations is below a predefined value EMaxrms and the local error at direct mapping (using transformation T) and the error of inverse mapping (using transformation T-1) of each one of the control points that were used to define the parameters are below a predefined value EMaxlocal. The method presents the following steps:

a. Set of active points: initially a set of active control points is defined, containing all the points.
b. Set of inactive points: initially an empty set of inactive control points is defined.
c. Transformation Tglobal and T-1global: parameters of an approximation of transformations Tglobal e T-1global are calculated using the control points which are in the group of active points.
d. Errors of Tglobal e T-1global: It is verified if the square average error for the transformations generated is below EMaxrms according to equation (4.2). It is also checked if the local errors of all points used for its definition are below EMaxlocal using equation (4.1) for both the direct mapping (using Tglobal) and inverse mapping (using T-1global). If all conditions are met, the transformations don't have outliers and the process is finished. Otherwise it follows to the next step.
e. Reordering of active points: The active control points are reordered from the point with highest error to that one with least error given by equation 5.3 which associates the original weight of the control point and the errors of direct and inverse mapping according to Tglobal and T-1global.
f. Transformations T1local e T1-1local: Starting from the active point of highest local error, one verifies if it is possible to obtain the parameters of valid transformations T1local and T1-1local using all the points of the group of active points, except the actual point. If it is not possible to calculate it, the next active point is tested. If the calculus was possible the next step is followed.
g. Errors of T1local and T1-1local: The errors of transformations T1local e T1-1local are verified in relation to EMaxrms and EMaxlocal. If any of the two conditions is not met, then the transformations T1local and T1-1local are discarded, and it goes back to step "f", considering the next active point ordered.
h. Reordering of inactive points: The list of inactive control points is reordered from the point of least local error to the largest one, associating the mapping errors of these points according to transformations T1local and T1-1local. The error calculus of each point is made according to equation 4.3.
i. Transformations T2local and T2-1local: Starting from the inactive point of least local error, it tries to calculate the parameters of transformations T2local e T2-1local using all points of the group from active points, including the inactive point.
j. Reutilization of the point: If possible the calculus of T2local and T2-1local and its errors are in accordance with EMaxrms and EMaxlocal then T1local gets the same parameters of T2local and T1-1local gets the same parameters of T2-1local. The inactive point used is moved to the list of active points.
k. Return to step "h" in case there are still inactive points to be tested.
l. The active point which was ignored at step "f" is moved to the list of inactive points.
m. If there is still any active point to be tested, the step "e" is executed again, using the next point.

$$E_{local}(Tp(p_1, p_2)) = EucDist(T(p_1), p_2) \quad (4.1)$$

Where *Tp* is a control point composed by a pair of coordinates in the form $p_1(x,y)$ and $p_2(u,v)$, according to a defined geometric transformation *T*, and *EucDist* is the Euclidean Distance between two points.

$$RMS = \sqrt{\frac{\sum_{k=1}^{n} {E_{local}}^2 (Tp_k)}{n}} \tag{4.2}$$

Where *n* is the number of control points (*Tp*) used to define the geometric transformation parameters.

$$Err\ (tp_k) = \frac{NDMap(tp_k) + NIMap(tp_k)}{2 \times TPWeight} \tag{4.3}$$

Where $NDMAP(tp_k)$ and $NIMAP(tp_k)$ represent respectively the direct and inverse mapping error, from the control point tp_k under a defined geometric transformation *T* normalized between [0,1] taking into account the direct mapping error of all control points. *TPWeight* is the original normalized control point weight which is one of the products of the step of generation of control points previously described.

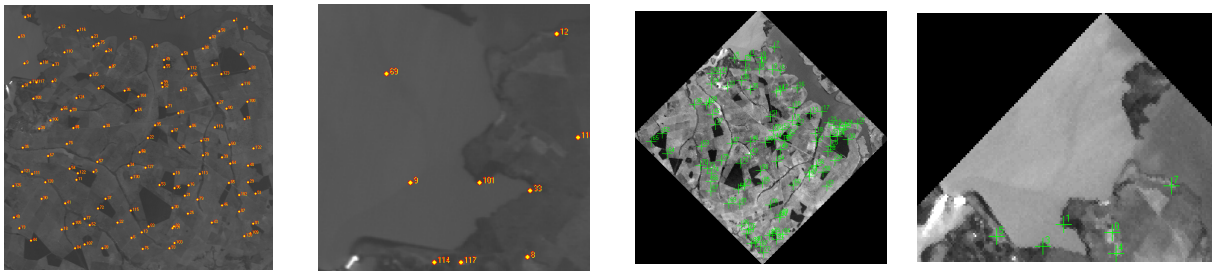
At the end of the process, if at least one transformation Tlocal, and the respective T-1local was generated, both of then satisfying the errors criteria EMaxrms and EMaxlocal, they become the final result. The generated transformations and the respective control points can be used by the final step to define geographic limits for the corrected image and the value of each image element.

5. Experiments and results

Some of the executed experiments results will be presented in this section. They were selected in which way to represent the major contributions of the proposed method and are meant to be easy reproducible as possible since they were generated with the help of a freely available prototype constructed in the form of a plugin for the system TerraPixel, TerraPixel, (2010), using the functionalities provided by library TerraLib, Câmara et al, (2008). The previews work from Fedorov, (2002) obtained comparison tests results by using the Regeemy prototype, Regeemy, (2010), also freely available for download.

5.1 Tie-points stability

For comparison effect between the new proposed method and its predecessor reference method two test images were used: a CBERS2 CCD scene crop (850x850 pixels, 20 meters of pixel size, band 2) and a second synthetic image based on a resampled version of the first one plus a contrast improvement and a 45 degrees rotation. The first pass was the tie-points generation using Regeemy (parameters: affine transformation, 512 initial interest points). 133 tie-points were generated and the resultant transformation has an RMSE error of 2.5868 pixel units with tie-points local errors values ranging from 0.0 to 7.0 pixel units. These values were taken as reference when running the proposed method and the maximum tie-point error parameter was set to 7 pixels units, affine transformation, and fully automatic interest point detection. 92 tie-points were generated and the resultant transformation has an RMSE error of 1.90 pixel units with tie-points local errors values ranging from 0.24 to 5.45. Like is shown in Figure 1, both transformations were calculated with good tip-points distribution all over the images, but the transformation generated by the proposed method has small errors and was obtained with a few tie-points taken over better image regions with more stable image features (like strong object corners) and avoiding homogeneous image areas (see figure b and d zooms) where weak unstable tie-points would appear.



(a) Reference method (b) Reference method Zoom (c) Proposed method (d) Proposed method zoom

Figure 1. First comparison results.

5.2 Geometric filter effectiveness

For this test the same pair of images will be used to check the effect of the geometric filter over the tie-points generated by homologous points matching stage. The first number of interest points is always kept to 500 (reference value taken from the reference method implemented by Regeemy). Without using the filter an unsuccessful result is obtained and the generated transformation has an RMSE error of 252.31 pixel units and most of the 141 generated tie-points are presenting local errors above 100 pixel units with a maximum tie-point local error of 712.42 pixel units. By enabling the geometric filter with affine transformation model and maximum tie-point local error restraint set to 10 pixel units the result is much better, as expected: 104 best tie-points are selected and the transformation has a RMSE error falls to 2.30 pixel units with local tie-points errors ranging from 0.18 to 7.51 pixel units.

To better analyze the proposed geometric filter behavior when matching those two images the following procedures were performed: Fixing the affine model with the maximum local tie-point error to 7.0 pixel units (reference value taken from the result from the topic above) and performing increments on the number of initial tie-points the result shows that the resultant transformation errors converge to a value of 1.78 pixel units when the number of initial tie-points is 512 as showed on Figure 2 . As expected the number of useful tie-points reaches its maximum value surrounding 512 for the test image. This behavior is also expected for larger images where the number of useful tie-points grows as the amount of image data gets large.

To analyze the filter behavior for different values of maximum allowed tie-point local error the number of initial tie-points was set to a fixed value of 512 and same test was performed. The result is showed at Figure 3, where the observed behavior is expected since for each maximum error increment the corresponding increment is also reflected on the global RMSE and number of tie-points.

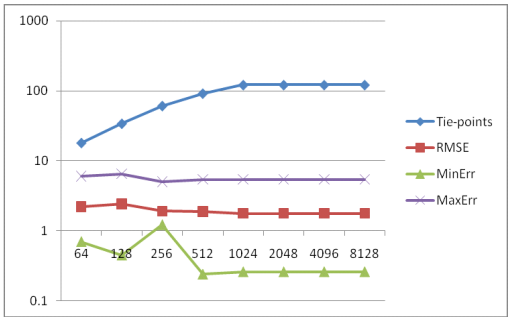


Figure 2. Filter behavior for multiple values of initial tie-points.

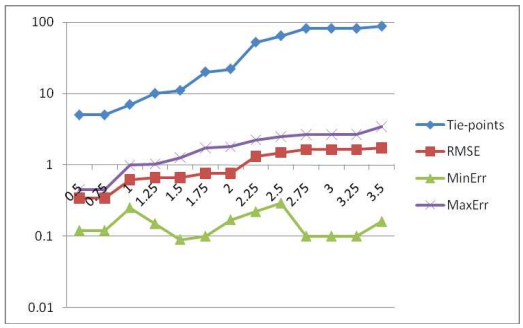


Figure 3. Filter behavior for multiple values of maximum local error.

5.3 Full scene automatic correction

For this experiment, a LANDSAT7/ETM image (orbit 222/073, acquisition date 06/09/2001, band 4, 30 m resolution) was used as a reference. The image to be adjusted is a scene from CBERS2/CCD (level 1 processing, orbit 159/121, acquisition date 03/30/2004, band 4, with 20 m of spatial resolution). Both images were obtained from the free images catalog from Brazil's National Institute for Space Research (INPE), CDSR, (2010). By using the affine geometric transformation model, 60m of maximum local tie-point error, 5000 initial tie-points the resultant geometric transformation was successfully generated with a RMSE error of 1.32 pixel units using 56 tie-points with local errors ranging from 0.31 to 3.33 pixel units. Note the CBERS scene covers only a portion of the LANDSAT image as showed in Figure 4 - Full scene matching result. With the generated parameters the adjust image can be successfully geometric and positioning corrected.

(a) LANDSAT (b) CBERS

Figure 4. Full scene matching result.

6. Conclusions

The main objective of this article is to present a complete and reproducible method of generic geometry correction highlighting outliers removing strategy. The main idea is based on previous work done by Fedorov, (2002) and goes a step forward by incorporating a new geometric filter approach used in conjunction with a more stable method to detect interest points. Since the main focus isn't the detection of interest points itself it can be replaced by any other stable method like the multi-scale approach presented in Fonseca, (1999).

According to the results the presented this method allows the effective correction of geometry and positioning when applied over satellite images from sensors with medium resolution. It also breaks through some restrictions when compared with the base work, where the matching was unable of give good results when there is a high amount of features to be matched like in the full scene case or because the less stable interest points approach.

The application of the method can also be done for higher resolution images, but to do that, an analysis is needed to adjust the parameters, which can be changed for that purpose (Moravec's window radius, correlation window radius, repeatability thresholds) similarly as succeeded for medium resolution images.

The method can also be easily modified, so it can be applied on data of different origin, such as radar data. For this finality, the generation of descriptors and its comparison can be adjusted for the use of new metrics and attributes, such as the use of common information, as suggested by Viola and Wells, (1997).

Although the method itself can be computationally expensive, it was verified with the implemented prototype that it is possible to make its execution faster, using procedures of code parallelization of the steps: image loading, image resampling, generation and matching

of descriptors.

References

1. H. P. Moravec. **Towards Automatic Visual Obstacle Avoidance**, Proceedings of the 5th IJCAI, MIT, Cambridge, Mass., 1977, p. 584.
2. Fedorov, D. **Sistema Semi-Automático de Registro e Mosaico de Imagens.** São José dos Campos: INPE, 2002. 150p. – (INPE-9582-TDI/838).
3. USGS Publications Warehouse. Extraction of GCP chips from GeoCover using Modified Moravec Interest Operator (MMIO) algorithm, **The United States Geological Survey (USGS) Online Repository**, available at <http://www.usgs.gov>.
4. Antonin Guttman. R-Trees: A Dynamic Index Structure for Spatial Searching. **ACM SIGMOD: International Conference on Management of Data**, 1984, pp. 47-57.
5. Martinuzzi S, Gould WA, Ramos González OM. 2006. **Creating cloud-free Landsat ETM+ data sets in tropical landscapes: cloud and cloud-shadow removal.** Gen. Tech. Rep. IITF-32. Río Piedras, PR: U.S.D.A., Forest Service, IITF. 12 p.
6. SILVA, W.A., IMAI, N.N.; POLIDORIO, A. M. Detecção de Nuvens e Sombras em Imagens CCD/CBERS por Correlação de Pontos Candidatos. **Simpósio Brasileiro de Sensoriamento Remoto**, Florianópolis, Brazil, 2007. p. 6199-6206.
7. Fonseca, L.M.G. **Registro Automático de Imagens de Sensoriamento Remoto basedo em Múltiplas Resoluções**. Doutorado em Computação Aplicada, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP, 1999.
8. Castejon, E.F., Forster, C.H.Q., Fonseca, L.M.G. M, E. Avaliação de métodos de casamento de imagens para mosaico de imagens orbitais. **XIV Simpósio Brasileiro de Sensoriamento Remoto**, Natal, Brazil, 2009, INPE, p. 6805-6812.
9. Câmara, G., Vinhas, L., Reis Ferreira, K., et, al., 2008. TerraLib: An open source GIS library for large-scale environmental and socio-economic application. In: Hall, G.B., Leahy, M.G. (Eds.), **Open Source Approaches in Spatial Data Handling**. Springer, Berlin, pp. 247–270.
10. TerraPixel 1.0.3b - An extensible image processing application prototype built using TerraLib classes and functions. Available at <http://www.dpi.inpe.br/terrapixel>.
11. Lowe, D.G., Object Recognition from Local Scale-Invariant Features. **International Conference of Compute Vision**, p.1150, Vol. 2, 1999.
12. Pratt, W.K. Correlation techniques of image registration. **IEEE Transactions on Aerospace and Eletronic Systems**. v.10 n. 3, p. 353-358, 1974.
13. Viola, P., Wells III, W. Alignment by Maximization of Mutual Information**. International Journal of Computer Vision**. Vol. 24, Num. 2, 1997.
14. Fischler, M.A., Bolles, R.C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. **Readings in computer vision: issues, problems, principles, and paradigms**, Morgan Kaufmann Publishers Inc., San Francisco, CA, 1987.
15. Catálogo de imagens - Instituto Nacional de Pesquisas Espaciais. Available at <http://www.dgi.inpe.br/CDSR>.
16. TerraView 3.3.1 – Um sistema de informações geográficas baseado em TerraLib. Available at <http://www.dpi.inpe.br/terraview>.
17. Umeyama, S., Least-Squares Estimation of Transformation Parameters Between Two Point Patterns. **IEEE Transactions on Pattern Analysis and Machine Intelligence**. Volume 13, Issue 4. Pages: 376 – 380. 1991. ISSN:0162-8828.
18. Fonseca L.M.G., Kenney C., Control point assessment for image registration. **XII Brazilian Symposium on Computer Graphics and Image Processing**. Campinas, SP, Brazil. Los Alamitos, CA: IEEE Computer Society, p. 125-132, 1999.
19. Regeemy – Automatic Image Registration and Mosaicking System. Available at <http://regima.dpi.inpe.br>.
20. Zitová, B. Flusser, J. Image registration methods: a survey. **Vision and Image Computing**, 21, 977-1000, 2003.
21. Fonseca, L.M.G., Manjunath, B.S. Registration techniches for multisensor remotely sensed images imagery. **Photogrammetric Engineering and Remote Sensing**. vol.62 no.9, 1049-1056. 1996.