AUTOMATIC REGISTRATION AND MOSAICKING SYSTEM FOR REMOTELY SENSED IMAGERY

Leila M. G. Fonseca¹
Dmitry Fedorov²
B. S. Manjunath²
Charles Kenney²
Emiliano Castejon¹
José Simeão de Medeiros¹

¹Instituto Nacional de Pesquisas Espaciais
Divisão de Processamento de Imagens
Av. dos Astronautas 1758, São José dos Campos, 12227-010
leila@dpi.inpe.br

²University of Santa Barbara, California
Dept. of Electrical and Computer Engineering
CA, USA
fedorov@ece.ucsb.edu

RESUMO

Registro de imagens é uma operação importante em várias aplicações em sensoriamento remoto e envolve, dentre outras operações, a identificação de pontos de controle correspondentes nas imagens. Como a identificação manual de pontos de controle é uma tarefa que toma muito tempo e também cansativa, várias técnicas automáticas têm sido desenvolvidas. Este trabalho descreve um sistema de registro e mosaic automático de imagens de sensoriamento remoto que está sendo desenvolvido pelo Instituto Nacional de Pesquisas Espaciais (INPE) em cooperação com a Universidade de Santa Bárbara, Califórnia (UCSB). O usuário pode interagir com o sistema fornecendo informações que são usadas para acelerar o processo de registro e também evitar erros no processo de casamento dos pontos de controle. Baseado em procedimentos estatísticos o sistema dá um indicativo da qualidade do registro, o que permite ao usuário parar o processo, modificar os parâmetros de registro ou aceitar os pontos de controle identificados. Vários testes com diferentes tipos de dados (imagens ópticas, radar, multisensóres, vídeo e imagens de alta resolução) têm sido realizados para verificar o desempenho do sistema. Um sistema demo está disponível na internet (<http://regima.dpi.inpe.br/>) e contém vários exemplos que podem ser executados na web.

Palavras chaves: registro de imagens, mosaico, pontos de controle.

ABSTRACT

Image registration is an important operation in many remote sensing applications and it, besides other tasks, involves the identification of corresponding control points in the images. As manual identification of control points may be time-consuming and tiring, several automatic techniques have been developed. This paper describes a system for automatic registration and mosaic of remote sensing images under development at The National Institute for Space Research (INPE) and at The University of California, Santa Barbara (UCSB). The user can provide information to the system in order to speed up the registration process as well as to avoid mismatched control points. Based on statistical procedure, the system gives an indication of the registration quality. This allows users to stop the processing, to modify the registration parameters or to continue the processing. Extensive system tests have been performed with different types of data (optical, radar, multi-sensor, high-resolution images and video sequences) in order to check the system performance. An online demo system is available on the internet (<http://regima.dpi.inpe.br/>) which contains several examples that can be carried out using web browser.

Key words: image registration, mosaic, control points.
1. INTRODUCTION

The increasing number of satellite images has reinforced the need for automatic image registration methods. Several techniques for automatic image registration have been developed (Fonseca and Manjunath, 1996; Zitová and Fluscher, 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). Besides, in many cases, it is impossible that the registration process be satisfactorily performed without human assistance. In order to overcome these problems we have developed an automatic system for image registration and mosaic which integrates many tools and information provided by the user in order to assist in the processing. This system has been developed by the Image Processing Division at INPE in cooperation with the Vision Lab at UCSB.

The developed registration system is a full-featured application intended for operational use by beginners as well as by advanced users. Registration may be achieved by one simple click or may be controlled by several parameters. The system contains toolboxes that increase the registration strength using the user’s knowledge. In addition, there is a pre-processing module that can change the image resolution, select a specific band, enhance histogram, etc. The control point extraction can be done within small windows in the images instead of using the whole image. This is very useful in cases of dense cloudiness, multi-temporal ocean shots, images with small overlapping areas, etc. Furthermore, the set of control points may be edited in a powerful embedded editor or exported to external applications.

Extensive algorithm tests have been performed by registering optical, radar, multi-sensor, high-resolution images and video sequences. In most cases the system has been successful. The main functionalities of the system will be described in detail in the next sections.

2. CONTROL POINT EXTRACTION ALGORITHMS

Three different algorithms for control point extraction have been implemented in the system. Moreover, the system has been designed so that other methods can be easily added. One of the algorithms uses optical flow ideas to extract the features from both images (Fonseca et al., 1999; Fonseca and Kenney, 1999). The second method uses the gray level information content of the images and their local wavelet transform modulus maxima to extract a set of control points (Fonseca and Costa, 1997; Fonseca, 1999). The latter uses centers of gravity of the closed boundaries and other strong edges as control points (Li et al., 1995).

The default registration method (optical flow) consists of a two-step procedure (Fig. 1). In the first step, candidate points in the first image are tentatively matched with candidate points in the second image. This preliminary matching is accomplished by matching features at each tie point in the first image with features at each tie point in the second image. These candidate control points are extracted using local maxima uniformly distributed on both images, processed with a variation of optical flow algorithm (Fonseca et al., 1999; Fonseca and Kenney, 1999).

![Fig. 1- Method based on optical flow.](image)

The user can select the number of control point candidates to start the algorithm with. A larger number increases quality but decreases speed exponentially. To overcome rotation effects, we have chosen windows about each feature point that have been rotated so that their central gradient points downward. This initial feature matching is used to eliminate points from both images that have no corresponding points. At the same time, an initial corresponding point set (control points) from the two images is obtained.

The second step refines the initial control point set by using a purely geometric matching procedure in which the location of points and their relationship to each other determine the matching. The geometric matching procedure is similar to the RANSAC method (Li et al., 1995) as described in Hartley and Zisserman (2000). Because the first and second steps are based on different matching criteria, they provide a safeguard against mismatched points. Moreover, the preliminary matching in the first step reduces the computational burden of the second step, resulting in a fast combined matching algorithm.

The second method is based on wavelet transform (Fonseca, 1999). Fig. 2 shows its flowchart. This method presents very good results for registering radar images. It is faster than the default method and it usually returns a larger number of control points. It’s
particularly useful for registering images of similar spectral bands (e.g. band 2 of SPOT with band 3 of Landsat TM). The feature points are extracted using wavelet transform local maxima. The algorithm initiates control points extraction process on the higher levels of wavelet distribution (lower resolution) and refines them on the lower levels (higher resolution). The point matching process is performed by maximizing normalized cross correlation over small windows surrounding the points. In order to remove mismatched control points a consistency checking procedure is performed recursively in such a way that the most likely incorrect match is deleted first, followed by the next most likely incorrect match and so on.

Contours are separated into two groups: closed and open. Each extracted contour from first image is compared against every contour from second image using shape attributes (perimeter, longest and shortest distances from the boundary to the centroid, etc) and correlation. The control points are identified as the centroids of the closed contours or salient points of the open contours.

2. TRANSFORMATION FUNCTION

Once the control points have been obtained, the transformation function is then estimated using the least squares method. At the present time three types of transformations are implemented: translation, similarity and affine (default). Another transformation of interest is the projective which has been implemented but still needs some adjustments.

Adequate transformation function selection is important to guarantee better registration quality. Suppose the real transformation between images is translation but there are some control points with bad quality in the set. In this case, undesirable small rotation can be avoided if the transformation selected is the translation.

Before determining the transformation, some consistency-checking procedure is performed in order to eliminate bad or mismatched control points. This is achieved using two independent tests. In the first one, a statistical procedure is used to characterize good and bad registrations (Kenney et al., 2003). The fit error is measured through mean and standard deviation of overlapping pixel differences. To characterize the bad and good registration, the fit error is computed using random values for the transform parameters. The fit error is then compared with both good and bad fit estimations in order to determine its acceptability. This “good fit-bad fit” statistical testing is supplemented with a bootstrap parameter variation test in which control point subsets are used to re-compute the transformation parameters. Large parameter variation over the subsets indicates the presence of mismatched control points. These consistency-checking tests provide a powerful mechanism to detect improper registrations. The second error detection procedure estimates the post-registration error between the first image and the back-transformation of the second image.

3. SYSTEM DESCRIPTION

The whole registration process is basically divided into three steps: image loading (File), control point extraction (Register), and image mosaic and registration (Mosaic). Fig. 4 shows the system interface. First, the images are loaded. In this step, the images can be pre-processed to improve their visual quality or adjust their resolution. The second phase is the control point extraction, which can be performed manually, automatically or semi-automatically by selecting rectangles in the images. The control points
can be edited so that those with larger errors (RMSE-Root Mean Square Error) can be removed from the list. The third step consists in generating an output image that can be a simply registered image or a mosaic.

The system has many tools to facilitate the registration process and improve the registered image quality. Some of them will be described below.

3.1 Image loading and pre-processing

Geographical parameters are loaded automatically for GeoTiff and GeoJP2 images. Other image formats might be geo-referenced by means of World File (.tfw), which defines pixel scale and geo-referencing and have to be read after loading a new image. Automatically retrieved pixel resolution is used to define scale difference between input images in order to facilitate the automatic control point identification task. If the images are not georeferenced and have different resolutions the operator must define the spatial resolution, in meters, for each image throughout pre-processing window. Afterwards, some pre-processing operations can be defined (Fig. 5).

The control point extraction algorithm uses images with 8 bits per pixel. Therefore any input RGB image will be seamlessly converted into 8 bits per pixel before control point acquisition. However, this transformed image will not be used to generate the resultant image but only to identify control points in both images. Standard conversion from RGB to gray scale image, which uses proportions of RGB channels, will be applied by default. It is advisable to select spectrally similar bands when multi-band or multi-sensor images are to be processed. Other pre-processing techniques could be used in order to improve the mosaic image quality as the light fall-off correction in aerial mosaicking.

3.2 Vignetting correction

Light fall-off and vignetting are common problems in aerial photographs. They are introduced by inevitable decrease of captured light by peripheral areas of the lens. Vignetting depends on the lens quality and its effect is not usually so noticeable. However, even a small effect can be a problem for a panorama when dark bands become evident. The amount of light fall-off is proportional to the cosine of the angle of view raised to the fourth power. Instead of using lens parameters to correct the problem, an approximated approach is used. Empirically, it was found that lightness increases proportionally to the distance from the center of the image raised to the second power and this hypothesis has been acceptable to correct those effects. Practically, the center is darkened and peripheral areas are lightened by half of the maximum correction amount in proportion to the distance from the center. The amount of maximum correction is identified by simple differences between mean values of the center and peripheral areas.

3.3 Automatic Control Point Extraction

Automatic control point extraction only needs one parameter, which specifies the registration quality. Basically, it is the number of control points to start with. This parameter is only modified in more complicated registration cases (images with clouds, severe temporal changes, very different sensors, images taken in side-look view). However, a reasonable number of control points (128 for default) provide satisfactory results. In the case of full scene images with small overlapping areas (less then 20%) it is advised to select similar geographical areas (rectangles smaller than 1000x1000 pixels) in both images, manually, to speed-up the search for control points. The operator may select more than
one rectangle in each image to improve the registration quality by iteratively adding new control points into a list. Fig. 6 shows the interface of the control points extraction process. The operator should select the type of control point extraction algorithm, the number of control points to start with, the type of transformation or to select rectangles if it is the case. Once all parameters are defined, the control point extraction process can be run.

Considering that the control point identification process takes a long time when the images are very big, the system automatically sets the maximum size of the images to 1000x1000 pixels. Therefore, larger images and the rectangles selected by users are down-sampled when their size exceeds this limit.

3.4 Geo-referencing strategy

In the system, the reference image defines the geographical position and projection for the resultant image. Geographical information is automatically extracted from GeoTiff and GeoJP2 images and may also be loaded from the World File. Resultant geographical position is recalculated during the image rendering procedure. The projection parameter is passed to the output image with no modifications. Therefore, it is up to the user to choose the right projection for the mosaic. The resultant geo-referenced image may be saved as GeoTiff as well as GeoJP2 and geographical parameters can be saved into a world file.

3.5 Mosaic rendering

The mosaic is actually an image rendering procedure (Fig. 7). The output image can be the mosaic itself or each one of the images can be rendered separately into universal (for both images) canvas or even a mosaic of both images into separate channels for matching verification.

The output image can be cropped into the reference image canvas or into the overlapping area. Image geometric transformation is obtained through common interpolation techniques: Nearest Neighbor or Bilinear. Besides, the images can be equalized using mean and standard deviation of the overlapping area.

The overlapping area of the mosaicking can be rendered either as the reference or warp image (normal rendering), interlacing both images or blending by weighted distances, which is very useful for mosaics of aerial photographs and video sequences.

Let \( G \) and \( g \) be the pixel values in the reference and warp image, respectively. Let \( (x,y) \) be the pixel position in the output image \( M \). Let \( dG \) and \( dg \) be the minimum distances between the pixel position \( (x,y) \) and border pixels of \( G \) and \( g \), respectively. Therefore, the new pixel values in the output image is calculated as:

\[
M(x,y) = \frac{dg/d}{dG/d} g(x,y) + \frac{dG/d}{dg/d} G(x,y),
\]

where \( d = dg + dG \).

Fig. 8 compares two mosaics obtained by normal and minimum distance blending algorithms. The arrow marks point out the bad rendering in the image.
borders. Another blending method based on minimum error has been under development.

Fig. 8 - Mosaic performed using (a) normal and (b) minimum distance blending algorithms.

3.6 Mosaic rendering with cloud removal

The cloud removal procedure is quite simple and therefore fast. It is executed over rendered resultant images considering information of the original images. For the entire overlapping area the difference between two input images is computed considering brightness of the channels. The cloud map is generated establishing that clouds gray level values are brighter than 100 and shadows gray level values are lower then 50. Having in mind that clouds and shadows are usually sufficiently big objects and that cloud contours are usually not well defined due to partial transparency of peripheral areas, the difference map is cleaned-up using simple morphological operators. Initially, erosion with 3x3 cross kernel is iteratively executed and is followed by a greater number of dilation iterations. This processing removes glare of cloud semi-transparent borders. Iterations number was established empirically. We have used four erosion and eight dilation iterations. The map is used to indicate the areas to be rendering to cover the clouds and shadows.

Fig. 9 shows two mosaics performed without (Fig. 9a) and with (Fig. 9b) cloud removal procedure.

4. IMPLEMENTATION CHARACTERISTICS

The system is developed for a multi-platform environment and is compiled and used under Solaris, Linux and Windows. It is written using common C++ and Trolltech Qt library (Trolltech, 2003). Most popular image formats (BMP, GIF, JPEG, PNG, RAW, TIFF, GeoTIFF, JPEG2000, GeoJP2, etc.) are accepted and mosaic results with geo-referenced base image can be saved as GeoTiff or GeoJP2 images. We have developed or modified several image file handlers for Qt, such as RAW reader/writer, JPEG encoder/decoder using Independent JPEG Group library, TIFF encoder/decoder using libTIFF from SGI, GeoTIFF encoder using libgeotiff and JPEG2000 and GeoJP2 using modified JasPer library.

In order to simplify the image viewing and handling operations, a universal viewer has been developed. It provides basic read/write and preview operations, such as: load file, save file, change zoom, image information, full screen preview, clipboard handling, etc. The visualization module also contains a layer to handle special information like control points, rectangles, etc.

Logically, the system is separated in two parts. The first part is the main driver for registration and mosaic. And the second is the GUI (Graphical User Interface) or command line module. The command line module is intended for calls from other systems; it’s used by the web demo. The GUI module is written in C++ using Trolltech Qt library (Trolltech, 2003).

The system can be used under several operating systems like Windows, UNIX variations and MAC OS. Implemented registration algorithms are quite fast to provide almost real time results on modern personal computers (P3 500MHz). Required memory depends on the image size. For full scenes such as Landsat TM/ETM+ (7000x7000 pixels), it is recommended a computer with 512 MB of RAM memory to avoid the excessive use of virtual memory. The amount of virtual memory is very important for the normal operation of the system; it is advisable to use between 400 and 2000 Mbytes in the case of large images.
5. SYSTEM OPERATION

The step-by-step example of a two image mosaic is presented in this section. However, only the basic features will be explained. For other features, the embedded software help system can be used.

Step One – Input image selection.

The system designates reference image as Image-1 and warp image as Image-2. To open the images, the Browse buttons must be used, indicated by numbers 1 and 2 (Fig. 10).

Fig. 10 – Registration Control.

Each image will be opened in its respective window (Fig. 11). The image window caption contains: zoom factor, image number, and file name, respectively.

Fig. 11 – Image 1 Preview Window.

A tool bar, indicated by number 3, can be used for image saving, pointer tool selection, zoom changing, and other operations. The image window also shows image size and color resolution, cursor pixel position, and tool information in the bottom status bar indicated by number 4.

Step 2 – Tie-point generation

In order to select matching areas, the button “Select Rectangles” must be used as indicated by number 5 (Fig. 12) on the “Register” tab. One rectangle must be selected from the reference image and another one from the warp image window, as indicated by numbers 9 (Fig. 13) and 10 (Fig. 14).

Control points can also be automatically extracted by skipping this pass and pressing “Auto Find Tie Points” button, indicated by number 6 (Fig. 12). If the list of control points isn’t empty, the system will ask about what should be done with the new control point set.

Fig. 12 – Tie-points selection.

Fig. 13 - Rectangle select (Reference Image).
Step Three – Tie-point edition

Control points extracted in the previous step can be seen in image windows as well as in a Tie Point editor window (Fig. 15), accessed by pressing the button indicated by number 8 (Fig. 12).

Once a point is selected the difference (in pixels) between reference and transformed control points is presented in the tool bar, indicated by number 11 in Fig. 15. The selected point is also shown on the image windows.

Tie points can be saved for later use. There are two supported formats: Regeemy internal format and ENVI point file format. Tie-points can be saved by using the “File → Save As” menu, as indicated by number 12 (Fig. 15).

Step Four – Mosaic rendering

To register or mosaic images, select Mosaic tab at the Registration Control window as indicated by number 13 in the Fig. 17.

The mosaic output depends on the selected rendering type (indicated by number 14 – Fig. 17), which can be selected in the following ways:
- Both Images: both images are adjusted and rendered in one RGB channel.
Different Channels: both images are adjusted and rendered in separated RGB channels, generating a RGB composition.

Only First Image: the first image is adjusted to the second one and rendered.

Only Second Image: the second image is adjusted to the first one and rendered.

Register Only: the second image is registered based only on the first one.

Subtraction: the rendered image contains a pixel-by-pixel normalized subtraction between the adjusted images.

Absolute Difference: each pixel of the rendered image will contain the absolute value of a pixel-by-pixel difference between the adjusted images.

Qualitative Difference: the rendered image will contain white pixels indicating detected differences between the adjusted images and black pixels where no difference was detected.

Quantitative Difference: the rendered image contains a pixel-by-pixel subtraction between the adjusted images plus a gain of 128.

To generate the mosaic image, press the “Register or Mosaic” button as indicated by number 16 in Fig. 17. A new image window (RES) will be created (Fig. 18). This image can be saved in the disk by using the menu “File → Save as” on the same window, indicated by number 17 (Fig. 18) or pressing the button “save the result image” (Fig. 17).

6. REGISTRATION EXAMPLES

We have included very difficult image registration examples in order to show the strengths and limitations of our approach. Several examples can be executed using web browser in the system’s online demo (Fedorov et al., 2005).

Fig. 19 shows a registered CBERS-2 image taking a Landsat image as reference in the registration operation.

Fig. 20 shows a mosaic of two adjacent CBERS-2 images. In this case the radiometric equalization was performed.

Fig. 21 shows the mosaic of two aerial images highly deteriorated by JPEG compression. The second image contains a simulated rotation effect with no scale differences. These images demonstrate the invariance of the registration system in relation to high rotation and noise. Registration was automatically performed in 875 ms.

Fig. 22 shows a mosaic of two full scene Landsat images of river Parnaiba taken at different dates. The registration was performed selecting two rectangles in both images; the processing took approximately three seconds.

Fig. 23 shows a mosaic of an aerial IR video sequence of 107 frames with 256x256 pixels taken from Los Angeles city. Mosaic was generated using default sequence method, restricting transformation to translation with no use of equalization and blending operations. The processing took about 8 seconds with steps of 8 frames.

Fig. 24 shows a mosaic of an aerial video images sequence of Santa Barbara airport, California. A mosaic was generated using default sequence method and projective transformation.

Fig. 25 shows a mosaic of an aerial video images sequence of Amazon forest, Brazil. Fig. 26 shows a mosaic of two adjacent IKONOS images. Finally, Fig. 27 shows a mosaic of four adjacent CBERS-2 IRMSS images.

Other experimental results with great variety of images from urban, forest and agricultural areas, different and similar sensors, and very high and very low resolutions were performed and in all cases were obtained very encouraging results.

7. CONCLUSION

This paper has presented an operational system for automatic image registration. Useful features have been included in the system to help users in the image
mosaic and registration processes. This system has been satisfactorily used in various important remote sensing applications such as Amazonia deforestation analysis and Ecological Economic Zoning as well in the radiographic image registration (Dotto et al., 2005).

In the future, we intend to improve the projective transformation, as well the registration algorithms, radiometric normalization and mosaic features by adding new blending procedures (Bagli and Fonseca, 2005, 2006).

8. ACKNOWLEDGEMENTS

This work was partially sponsored by CAPES, SELPER (Brazil), Office of Naval Research and China Lake Naval Air Warfare Center and CalTrans (USA). The authors would like to thank Rodrigo Rizzi (INPE) for his helpful suggestions, which have enhanced the system features from the user’s point of view and for operational testing and important remarks. GeoJP2™ is a trademark of Mapping Science, Inc. The GeoJP2™ format is the intellectual property of Mapping Science, Inc.

(a)  
(b)  
(c)  
Fig. 19 - Image in (c) is a registered CBERS2 Level 1 image (b) superimposed on a LandSat image (a) used as reference.

(a)  
(b)  
(c)  
Fig. 20 - Image in (c) is a mosaic of two adjacent CBERS-2 images (a) and (b). The radiometric equalization operation was performed.

(a)  
(b)  
(c)  
Fig. 21 - Image in (c) is a mosaic of two urban aerial images (a) 306x386 and (b) 335x472 pixels.
Fig. 22 - Mosaic of two images (6607*5999 and 6628*5999 pixels) of Landsat TM-5.

Fig. 23 - Mosaic of an aerial IR video sequence mosaic: 107 images, 256x256 pixels, generated in 20 seconds.

Fig. 24 - Mosaic of a video image sequence of the Santa Barbara airport, California.
Fig. 25 - Mosaic of a video images sequence of Amazon forest, Brazil: 1100 images, 720x40 pixels, generated in 5 minutes.

Fig. 26 - Mosaic of two images of IKONOS.

Fig. 27 - Mosaic of 4 CBERS-2 IRMSS images taken in different times.
REFERENCES


