Automatic Fusion of SAR and Optical Imagery based on Line Features

Jan Dirk Wegner, Institute for Photogrammetry and Geoinformation (IPI), Leibniz University Hanover, Germany
Jordi Inglada, Céline Tison, Centre National d'Etudes Spatiales (CNES), France

Abstract

A generic image processing chain for the fusion of very high resolution optical and SAR imagery is developed. It is fast to operate and provides reasonably good accuracy. This method is capable of rapidly registering optical and SAR images based on extracted features rather than on pixels. Indeed, at very high resolution, pixel approaches fail on above ground structures. An approach on a higher semantic level is required. The fusion is generally applicable to many kinds of remote sensing imagery due to it being generic. Its goal is to enable feature based registration of unrectified optical and SAR images with sub metric resolution in urban areas. Important aspects of this fusion approach are the very high resolution of the images, the introduction of a classification, rapid processing and the modular construction of the entire processing chain based on the open source software library ORFEO Toolbox (OTB)[1].

1 Introduction

Today, images from a variety of high resolution optical and SAR sensors are available. They are either airborne (Ramses, Pelican, etc.) or space borne (TerraSARX, CosmoSkyMed, etc.). New application possibilities arise due to sub metric resolution and the complementary properties of optical and SAR sensors. For example, in the field of rapid change detection after natural hazards, highly automated image analysis of combined optical and SAR imagery is necessary. Therefore, a highly automated generic optical/SAR fusion method was developed that is based on line features. It is fast to operate and has a component-wise architecture thus allowing for its adaptation to particular scenes. Additionally, it applies rather general sensor models making it applicable to multiple kinds of sensors. It is based on previous work conducted at CNES [1] and ENST Paris [2].

The developed strategy is tested on simulated data with sub metric ground resolution. A set of optical and SAR aerial images is used (see **Figure 1**). The optical image was taken by the French National Geographic Institute (IGN). The SAR image was taken in X-Band by the aerial sensor RAMSES. It was provided by courtesy of the French DGA (Defence Procurement Agency).

The algorithm proposed in this paper was implemented in C++ as a part of the open source software library ORFEO Toolbox (OTB). It is part of the ORFEO Accompaniment Programme which was set up in order to prepare for the analysis of high resolution images, derived from the new sensors introduced by ORFEO. ORFEO is a French-Italian high resolution Earth observation satellite program, incorporating the optical system Pléiades (France) and the SAR system CosmoSkyMed (Italy). The approach presented in this paper relies on already developed algorithms of OTB and adds new tools. The proposed image fusion algorithm will be integrated into OTB after severe testing.

Figure 1 **Optical (left) and SAR (right) test images**

2 Image Fusion Strategy

2.1 Strategy Outline

The first fusion step consists of an ortho-rectification of both optical and SAR image. Both images are projected from sensor space to object space. An external digital elevation model (DEM) is used in order to rectify distortions introduced by the terrain. Objects included in the DEM are rectified whereas objects which are not included stay distorted, in particular

[1] See the OTB webpage *http://otb.cnes.fr*

buildings. The transformations also account for sensor specific residuals. The second fusion step applies appropriate pre-processing filters to the rectified images. This component reduces the noise level and prepares the images for further analysis. In the following component, both images are classified. The classification distinguishes between rectified ground regions and unrectified buildings. This step is necessary because only rectified areas of the images can be considered for the registration of the images.
After all rectified image regions have been identified, features are extracted. The outcome of this fourth registration step is two images displaying extracted lines. Thereafter, distance images are calculated from the feature images. They are registered using the Insight Toolkit[2] (ITK) registration framework [3]. After convergence of the registration towards an optimal solution, the final registration parameters are used to fuse the SAR image with the optical image.

2.2 Ortho-Rectification

Modelling has to be conducted in order to account for the different viewing geometries of SAR and optical sensors (see **Figure 2**). Distortions due to the terrain have to be taken care of, too. Both images are projected from sensor space to the ground as a first geometric registration step. A general a priori model-based approach was chosen, capable of rectifying the images without in depth knowledge of sensor parameters. An external DEM is used to reduce distortions introduced by rough terrain.

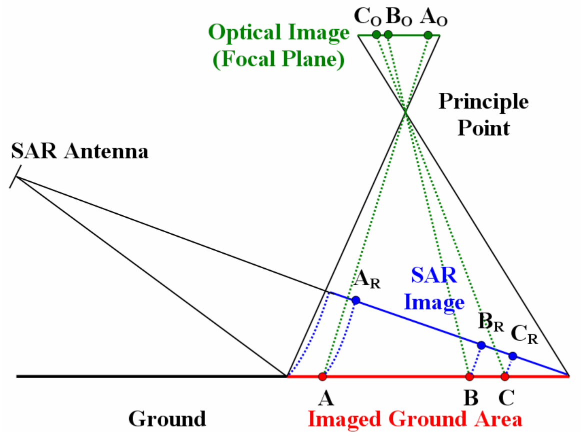

Figure 2 **Optical and SAR sensor geometries**

2.2.1 Optical Sensor Model

For optical imagery, the inverse 3D collinearity equations (object to image) are used in order to project the image to the ground. An indirect geometric image transformation for each pixel of the ortho-image is conducted. The pixel size of the ortho-image is selected corresponding to the ground resolution of the sensor. For all raster points of the ortho-image on the ground, the corresponding height values have to be interpolated within the DEM. The entire geometric modelling process is conducted in physical coordinates. The interpolation of the grey value within the original image in sensor geometry is a simple bilinear interpolation.

[2] See the ITK webpage http://*www.itk.org* for details

2.2.2 SAR Sensor Model

The SAR image is projected to the ground with the inverse equations proposed by Toutin [4], originally derived from the collinearity equations. They incorporate three different models: the motion model, the sensor model and the earth model. Hence, three coordinate systems are used: the image coordinate system, the intermediate coordinate system and the ground cartographic coordinate system. The first step is a transformation of the ground coordinates to the intermediate coordinate system. It simply applies one translation and one rotation. Furthermore, the coordinates of the intermediate system are transformed to the image coordinates.

2.3 Pre-processing

For the optical image, an anisotropic diffusion filter already existing in OTB was chosen, in order to preserve edges. It implements an N-dimensional version of the anisotropic diffusion equation for scalar-valued images proposed in [5].
In SAR imagery, both noise reduction and the treatment of the speckle effect have to be taken into account. For good line detection results, the preservation of edges by the anti-speckling filter is imperative. A filter has to be chosen that restricts smoothing to rather homogenous regions. Local statistics (e.g. mean, standard deviation) and texture parameters (e.g. contrast, entropy) can be calculated within the filter matrix for deciding whether a pixel belongs to a region or not. Hence, edge preserving speckle reduction was conducted with the Frost filter [6].

3 Classification

The main reason for the introduction of a classification is that the image fusion has to take place in rectified regions of the images. Objects present in the image but not in the DEM have not been accounted for. In particular, buildings are not contained within a DEM, i.e. buildings in the images stay distorted. Hence, we have to distinguish between buildings and the ground in order to register the images only on the rectified ground level. Registering the images in building regions could immediately lead to severe perturbations. Hence, the main goal is to classify the image

into the rectified ground level and unrectified objects above ground.
We chose a classification based on Support Vector Machines because it is applicable to both optical and SAR imagery. It is a supervised learning method, allowing the operator to specify training regions for each class of interest. The optical image was classified into five classes, the SAR image into six classes (see **Figure 3**). The classes for the optical image are vegetation (green), ground (brown), roof (blue), façade (white) and shadow (black). SAR classes are vegetation (green), ground (brown), light roof (blue), dark roof (yellow), strong reflectors (red) and occluded areas (black). Several texture parameters and statistics were input to establish the SVM feature vectors, e.g. kurtosis and skewness. However, they did not significantly improve the results obtained with a simple mean image as input.

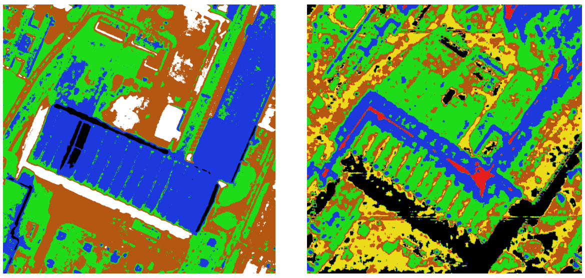

Figure 3 **Optical image (left) and SAR image (right) classified with SVM**

Apparently, the SVM classification results are not sufficient. They lack accuracy, in particular the classification of the SAR image. Hence, such classification results were not considered for the later on image fusion. Future works will have to comprise a refinement of the classification in order to enable full integration into the fusion process. A much better segmentation of the SAR image can be achieved with a classification based on Markov Random Fields (MRF). Previous work conducted in [7] has shown that a MRF classification assuming a Fisher distribution of the image leads to enhanced results. SVM classification results could be used for initializing the MRF classification.

4 Feature Extraction

4.1 Line Extraction

Regions, points and lines could be used in order to merge the optical and the SAR image. In our case, line feature proved to deliver the best results. In the optical image, lines were extracted with the Canny-operator [8]. An asymmetric fusion of lines approach was chosen for the line extraction in SAR imagery. It was originally developed for the detection of road networks [9]. Its strategy is the fusion of the outcome of two separate line detectors D1 and D2. Usually, this line extraction operator should be applied to the original Single Look Complex (SLC) image because it is already adapted to SAR imagery. However, line extraction was tested on both the SLC image and the Frost filtered image. For both cases we obtained good results. None the less, the outcome from the line detection algorithm applied to the Frost filtered image showed less small line pieces and was thus preferred for further processing.

4.2 Distance Mapping

Distance images are calculated from the feature images using the Danielsson approach [10]. This step was introduced in order to reduce remaining geometric residuals that result in different absolute positions of the extracted lines. The later on image fusion is more stable and converges faster if we use distance images instead of the feature images themselves. The coloured boxes frame corresponding parts of the images.

5 Fusion

5.1 Fusion Framework

The image registration process is embedded into a registration framework, originally provided by ITK. One input image is called the fixed image (it acts as the reference) and the other one is called the moving image. The goal is to find the optimal spatial mapping parameters that align the moving image with the fixed image. Hence, the moving image has to be deformed.
The registration framework treats image registration as an iterative optimization problem and consists of four components:

- a transform component applies a geometric transformation to the fixed image points in physical space in order to map them to the moving image,
- an interpolator evaluates intensities in the moving image at non-grid positions,
- a metric measures the similarity between the deformed moving image and the fixed image,
- and an optimizer optimizes the similarity value.

A new set of parameters for the transformation of the following iteration is determined after each optimization step. The major advantage of this modular conception of the registration framework is easy compatibility of a large variety of optimizers, geometric transformations and similarity measure techniques.

5.2 Results

Several test programs with a variety of component combinations were tried (e.g. mutual information, genetic optimizers). However, a combination of the simplest components was found to be the most robust. A metric based on normalized cross-correlation in combination with a regular step gradient descent optimizer led to the fastest convergence and the best results (see **Figure 4**). For testing purposes, a fairly simple transform was used, consisting of a centred rotation followed by translations in x and y. Bilinear interpolation was deployed for the computation of the grey values.

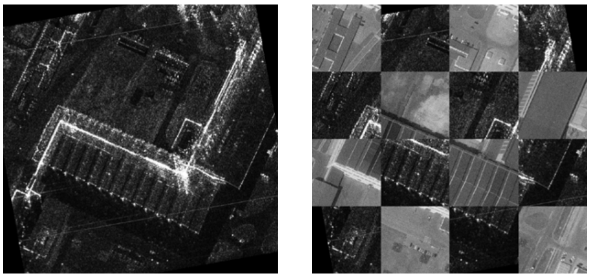

Figure 4 **SAR image rotated and translated (left) and fused with the optical image (right)**

6 Conclusion and Outlook

In conclusion, the overall registration strategy proved to be successful even without including the classification results. It is generic and relatively fast to compute. None the less, further refinement of all components has to be conducted. The ortho-rectification of the SAR image could possibly be enhanced with an approach based on Doppler and range equations [11, 12]. In order to improve classification results of the SAR image, a higher level classification has to be applied. For urban scenes, MRF classification techniques based on a modified Fisher distribution have been successfully tested [7]. As soon as classification results are used for the fusion, the fusion accuracy is considered to improve significantly. Additionally, an InSAR DSM will be integrated in order to enhance the classification. A more sophisticated transform for the ITK mapping component is necessary, too. At least, anisotropic scaling has to be integrated. Furthermore, non-rigid transformations should be tested. An algorithm based on the Finite Elements Method (FEM) already exists in OTB. It will be applied to the test images leading to disparity maps.

Acknowledgment

The authors thank ONERA and DGA for providing the test images. They also thank CS SI, in particular Thomas Feuvrier, for their support and fruitful discussions.

References

[1] Inglada, J.; Giros, A.: *On the Possibility of Automatic Multisensor Image Registration*. IEEE Transactions on Geoscience and Remote Sensing, 42(10): 2104 - 2120, October 2004.

[2] Galland, F.; Tupin, F.; Nicolas, J.-M.; Roux, M.: *Registering of Synthetic Aperture Radar and optical data*. In: IEEE International Geoscience and Remote Sensing Symposium, IGARSS '05, 5: 3513 - 3516, July 2005.

[3] Ibánez, L.; Schroeder, W.; Ng, L.; Cates, J.: *The ITK Software Guide.* 2nd edition, updated for ITK version 2.4, November 2005.

[4] Toutin, T.; Carbonneau, Y.; St-Laurent, L.: *An Integrated Method to Rectify Airborne Radar Imagery using DEM.* Photogrammetric Engineering and Remote Sensing, 58(4): 417 - 422, April 1992.

[5] Perona, P.; Malik, J.: *Scale-space and edge detection using anisotropic diffusion.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(7): 629 - 639, July 1990.

[6] Frost, V.S.; Abbott Stiles, J.; Shanmugan, K.S.; Holtzman, J.C.: *A Model for Radar Images and Its Application to Adaptive Digital Filtering of Multiplicative Noise*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 4(2): 157 - 166, March 1982.

[7] Tison, C.; Nicolas, J.-M.; Tupin, F.; Maître, H.: *A new statistical model of urban areas in high resolution SAR images for Markovian Segmentation.* IEEE Transactions on Geoscience and Remote Sensing (IGARSS'03 special issue), 42(10): 2046 - 2057, October 2004.

[8] Canny, J.: *A Computational Approach to Edge Detection.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(6): 679 - 698, November 1986.

[9] Tupin, F.; Maître, H.; Mangin, J.-M.; Nicoals, J.-M.; Pechersky, E.: *Detection of Linear Features in SAR Images: Application to Road Network Extraction.* IEEE Transactions on Geoscience and Remote Sensing, 36(2): 434 - 453, March 1998.

[10] Danielsson, P.E.: *Euclidean distance mapping.* Computer Graphics and Image Processing, 14: 227 - 248, 1980.

[11] Gelautz, M.; Frick, H., Raggam, J.; Burgstaller, J.; Leberl, F.: *SAR Image Simulation and Analysis of Alpine Terrain.* ISPRS Journal of Photogrammetry and Remote Sensing, 53: 17 - 38, 1998.

[12] Sörgel, U.: *Iterative Verfahren zur Detektion und Rekonstruktion von Gebäuden in SAR- und InSAR-Daten*. PhD Thesis, Fachbereich Elektrotechnik und Informationstechnik der Universität Hannover, March 2003.