# ON THE USE OF SAR AND OPTICAL IMAGES COMBINATION FOR SCENE INTERPRETATION.

G. Oller [a], D. Petit [a], J. Inglada [b]

[a] MAGELLIUM, 10 avenue de l'Europe, 31520 Ramonville St-Agne, France – guillaume.oller@magellium.fr
[b] CNES, 18 av. E. Belin, 31055 Toulouse Cedex 4, France

**Commission I, WG I/3 - MULTI-PLATFORM SENSING AND SENSOR NETWORKS**

**KEY WORDS**: SAR, image, Registration, Matching, Similarity Measures, Correlation, Data Fusion.

**ABSTRACT:**
Scene interpretation using multi sensors images becomes an active field of interest in Earth observation domain. In this context, the joint use of radar and optical images remains a hard task. Indeed, major differences exist between images, as well on geometric as radiometric features. But these differences also generate a high degree of complementarity between the two imaging systems.
This paper deals with opportunities offered by the combination of these imagery types, and with the difficulties to implement them. We first focus on the problem of automatic SAR and optical image registration, which is a preliminary and essential step for any kind of application (cartography, disaster monitoring, image interpretation…). Then, we present data fusion methods dedicated to SAR and optical image processing, which permits to improve image interpretation.
The first part of our paper addresses the crucial and difficult problem of SAR and Optical image registration. A fine image registration needs precise geometric modeling and auxiliary data (DEM, Ground Control Points…), which are not always available.
To solve this problem, the need of automatic registration techniques increases. Many algorithms are available, each one needing to define a measure of image similarity. Classical correlation coefficient relies on the assumption of a linear relationship between variables, which is not valid in the case of SAR and optical registration. Thus, we describe and evaluate methods based on statistical properties of signals. These methods are well known in medical imaging, and usually named "similarity measures" (Mutual Information, Cluster Reward Algorithm…). In this section, we also propose to compare statistical measures and more classical edge-based methods.
After the definition of a similarity measure, a matching strategy has to be developed. Our experimentations are mainly based on adapted gradient descent algorithms, which give good results in terms of robustness and performances. We describe these strategies, and evaluate registration performances at several resolutions.
The last part of our paper deals with using jointly SAR and optical image for scene interpretation. Radiometric differences make classical fusion methods not applicable. Geometric distortions also perturb fusion algorithms, because of non superposition of pixels. We describe a product developed for image interpretation, which includes specific tools adapted to SAR and optical fusion. We propose several fusion methods (wavelet based method for example) but also some fine local registration algorithms dedicated to SAR and optical image analysis and interpretation.

## 1. INTRODUCTION

With the increasing amount of data in the field of Earth Observation, multi-sensor scene analysis became an active field of interest. Complementarity of imagery systems and spectral bands allows development of several multimodal applications. In this context, the joint use of radar and optical images remains a hard task. Indeed, major differences exist between images, as well on geometric as radiometric features. But these differences also generate high degree of complementarity between both imaging systems.

Image registration is a preliminary and essential step for multimodal applications. But a fine image registration needs precise geometric modeling and auxiliary data (DEM, Ground Control Points…), which are not always available or sufficient. Thus automatic registration methods are considered in the first part of our paper. We focus on the difficulties to co-register SAR and optical images, on the necessity to use specific similarity metrics, and on the particularities of high or very high resolution imagery. Indeed, terrain elevation has important influence on registration, due to parallax between acquisitions. In section 3, a processing chain providing a "cartographic" registration (i.e. images registered on ground, but not necessary on buildings) is presented and evaluated.

Once images have been registered, scene interpretation combining layers is possible. In the last part, we present a trail for a fine coregistration / scene interpretation with metric resolution images. Then we conclude on the use of SAR and optical images combination for scene interpretation.

## 2. SAR AND OPTICAL IMAGES REGISTRATION

### 2.1 Decametric registration

Automatic registration techniques are based on a simple algorithm: an optimization process attempt to find the best parameters for a geometrical deformation model in order to match extracted features (from a first image) with simulated features (from a second image) according to the deformation model (Figure 1).
Features can be explicitly radiometry, but also edges, regions… A strategy improves optimization efficiency in complex case. The key concept of any matching algorithm is the similarity measure. Given a couple of images or set of features, its purpose is to estimate if the first element "looks like" to the second or not (according to a criteria). In the field of image processing, the normalized correlation coefficient is usually employed, but

it is not applicable in multimodal cases because it presumes linearity between variables.

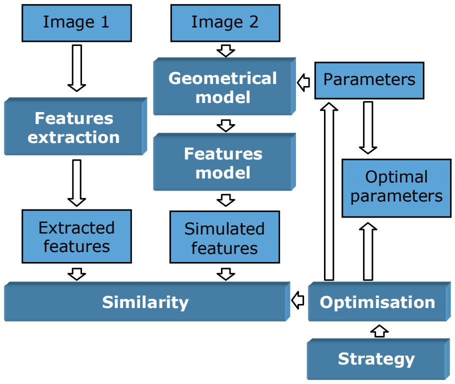



Figure 1. Scheme of a generic registration technique.

Nevertheless, the possibility of automatic multisensor image registration has been highlighted by recent works for different couple of sensors: for examples, SPOT5 and ERS [1], SPOT5 and ENVISAT (ASAR) [2], LANDSAT TM and IRS PAN [3], IKONOS and LANDSAT-7 [4], SEASAT and SPOT [5], RADARSAT and optical sensor [6].
Several techniques are based on statistical dependency measures like mutual information, or lesser sensor dependent features like edges. Mutual Information (MI) measure has been introduced in satellite image processing a few years ago [6, 7] rather than it is used in medical imaging for a long time [8].
These studies call attention to the aptitude of such matching technique to reach a sub-pixel registration accuracy for resolutions close to a decameter.

### 2.2 Metric registration

The achievement of high resolution imaging systems raises new problems for image registration. Indeed, geometric differences, due to acquiring technology, increase. This is particularly true in urban area, where the parallax effect is important on buildings, and trees (i.e. above-ground objects). Shadows, which are totally incoherent between each acquisition (especially between SAR and optical), also have an important influence on registration accuracy.
It implies that specific methods are required for a metric registration with multisensor images. Automatic [9] and manual methods [10] involves a minimal scene interpretation in order to attain a satisfying accuracy.
The next example (Figure 2) shows the influence of parallax and shadow on registration: A similarity surface is obtained by computing mutual information [11] for each translation of the slave image on the reference one. Image resolution is 1 square meter.
The influence of ground points and above-ground ones is clearly visible on the similarity surface: two maxima are present, one corresponding to the ground point's contribution, the other to the above-ground points. In this case, shadows effects are merged with building influence.

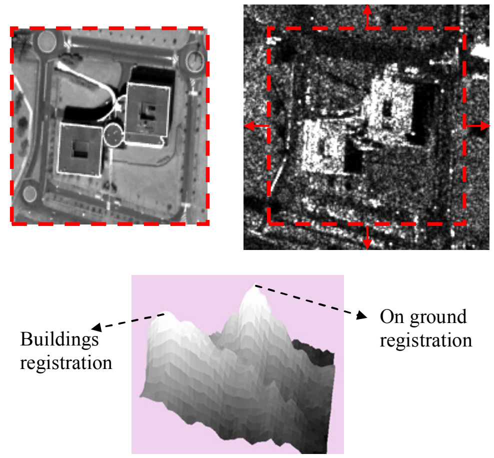
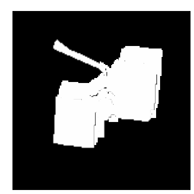



Figure 2. Surface of similarity: effects of buildings and shadows on registration

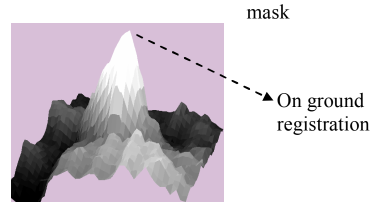
This simple example shows the necessity to integrate exogenous information for high registration accuracy.
A manual classification has been done for the previous case, in order to differentiate ground and above-ground features (Figure 3). Thus, above-ground points are not taken into account when estimating the joint probability for mutual information.

Looking at the similarity surface (Figure 3.c), accuracy is obviously increased. The second maximum disappears, and the main peak is thinner.

(a) Original image



(b) Above-ground & shadows mask




(c) Similarity surface using mask

Figure 3. Results using a ground/above-ground classification

## 3. HIGH RESOLUTION IMAGES REGISTRATION

### 3.1 Ground / Above-Ground Classification

Unless we intend to reconstruct a three dimensional scene, we ought to choose whether we aim to register ground pixel or above-ground pixels. We decide to get the best registration accuracy for ground pixels.
In the framework of a fully automatic approach, we propose in this section a simple and automatic method to classify urban and natural areas. In this case, not all the above-ground components are detected: vegetation is not taken into account. But urban classification permits to eliminate man made

structures which are mainly part of above-ground. The problem of vegetation detection will be treated in further works.

The method we used is based on a simple directional variances computation [12]. The starting assumption is that in urban areas, the variance of the signal is high in all directions. Thus, for each pixel, our detector computes the variance in 8 directions and returns the minimum value. If the returned value is high, the pixel is then considered to belong to an urban area. Figure 4 illustrates this procedure.

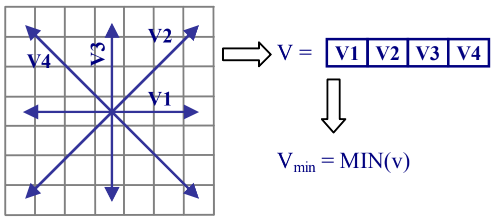



Figure 4. Illustration of the directional variances computation for 4 directions.

For higher processing speed, this detection is performed on decreased resolution images. Then, this information is regularized and used to exclude these points from Mutual Information computation in the registration step.
Note that this method can be applied either on optical or on SAR images (speckle is filtered by a multi-resolution scheme). Thus, the final mask is constructed combining both results.
A result of this classification is shown on Figure 5. The mask obtained, even if it is not of high accuracy, gives a rough estimation of buildings presence. In this first step, we only expect to improve registration accuracy by using a better ratio of ground pixel selection than an overall selection (this goal seems reachable). Further work will deal with classification quality improvement.

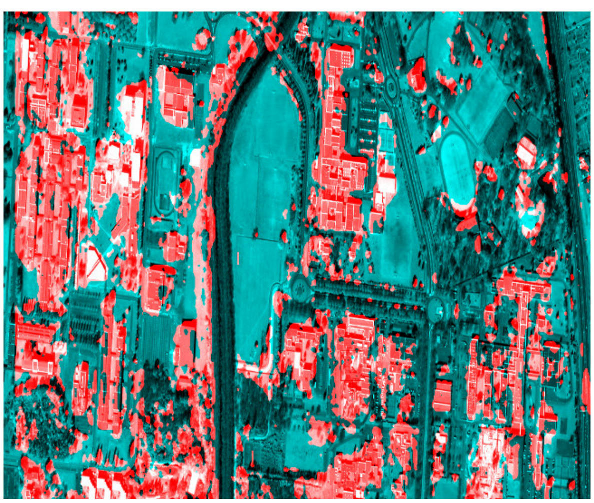


Figure 5. Results of urban area classification. Detected urban areas are in red.

### 3.2 Processing Chain

In this section the whole processing chain implemented to register high resolution SAR and optical images is presented. We recall that the objective is to generate "on ground" registration (i.e. cartographic).

The chain is composed by 4 steps: pre registration, ground/above-ground classification, image matching and final registration.

#### 3.2.1 *Pre-registration*
The aim of the pre-registration step is to roughly register images. Working with under-sampled data, an automatic process is performed to find the best rotation and translation between images according Mutual Information. This preliminary step can be avoided if the geometrical model and its *a priori* parameters give a satisfactory localization.

Figure 6. Results of the pre-registration step.

#### 3.2.2 *Classification and Matching*
For pre-registered images, the classification method described in section 3 is applied. Two masks are obtained and combined to form a final urban classification. As explained in section 2.2, pixels belonging to this mask will not be considered for image registration.
Matching is performed on a regular grid. For each point, its homologous in the second image is searched, considering only translations (this is possible because a pre-registration has been done). Points of the grid included in the urban mask are not processed. For each window, Mutual Information is estimated only on non-urban points. A few hundreds of tie points are generated.

#### 3.2.3 *Final Registration*
Final registration is performed by estimating the best deformation model fitting the tie points collected. The objective being to register images on ground, we assume that the model we search is low frequency (2$^{nd}$ order polynomial for example).
In a preliminary step, tie points have to be filtered looking at their similarity value. RANSAC algorithm [13] can also be used to reject outliers' points. To estimate the final model via a least square method, only a few tens of the best matches are preserved. Figure 7 shows samples of registration results. Some points of interest located on ground and their homologous have been identified with colored lines in both images.

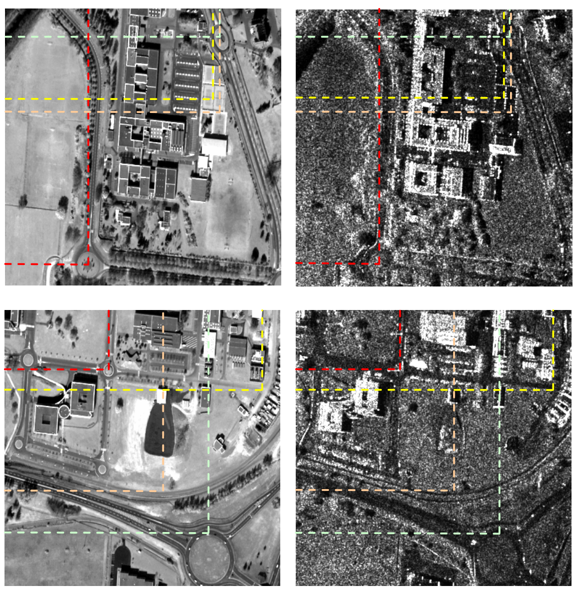
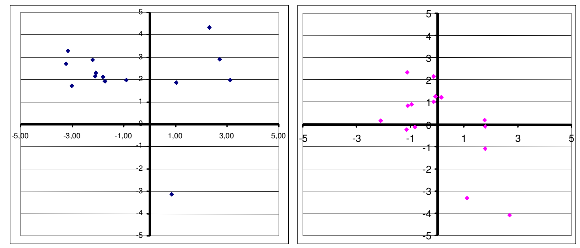


Figure 7. Parts of the registered optical and SAR images. Roads, roundabout and grass areas are correctly registered whereas buildings not.

### 3.3 Results

The main objective of our work was to evaluate the improvement of registration when integrating a classification in the process. With this intention, we quantified on-ground registration by manually selecting some interest points on ground in both images.

Error between automatic and manual registration is presented on Figure 8. The improvement of registration accuracy using a classification mask is clear (Figure 8.b). We observe a reduction of parallax effect (due to above-ground structures).



(a) (b)

Figure 8. Registration error according to x axis and y axis, evaluated on manual control points, with (a) standard registration method, and with (b) registration integrating a non-urban mask.

| | Registration without masking | Registration with masking |
|---|---|---|
| RMSE in x | 2.6 m | 1.4 m |
| RMSE in y | 2.5 m | 1.5 m |
| RMSE | 3.6 m | 2.1 m |

Table 1. Registration accuracy computed with 15 control points.

## 4. SCENE INTERPRETATION

A fine co-registration allows an eventual data fusion and a scene interpretation. Nevertheless, it requires a precision better than one pixel difficult to achieve with standard methods and DEM. Thus, three approaches are usually considered:

- manual registration method performed by a specialist of image interpretation, who can be assisted by image analysis tools;
- refine co-registration which leads to a production of a Digital Elevation Model;
- scene interpretation fusion where scene interpretation are individually done for each image and only the results are co-registered and fused.

The first method is largely practiced nowadays, and some new concepts of interactive tools and human machine interface have been tested in a Magellium's software (MAGFUSION).
The second principle is the basic concept of photogrammetry or radargrammetry but stereo possibilities with very different sensors are still under studies.
The third method consists in proceeding to a scene interpretation before the final registration, which permits to process each image with adequate tools. Unfortunately, this method does not benefit from the multisensor approach for scene interpretation.
Other methods can perform fine registration and scene interpretation iteratively. But, in fact, they all need to find optimal parameters for both the geometrical model and the features model (see figure below). Parameters of features model depend on sensor and scene contents; they therefore define the scene interpretation.

The best process could be to optimize both parameters, but computation times are generally discouraging. Indeed, scene content description often needs many parameters.
However, we demonstrated in previous part, that we do not need an accurate scene interpretation model to improve registration accuracy. It means that adding rough models (for buildings and shadows for example) to the optimization process can be a realistic solution.

This method could take profit of the knowledge of the scene contents (or the presupposed contents):

- to improve the use of the geometrical model;

- but also to use more appropriate similarity measures between the too much constrained correlation and the few constrained mutual information.

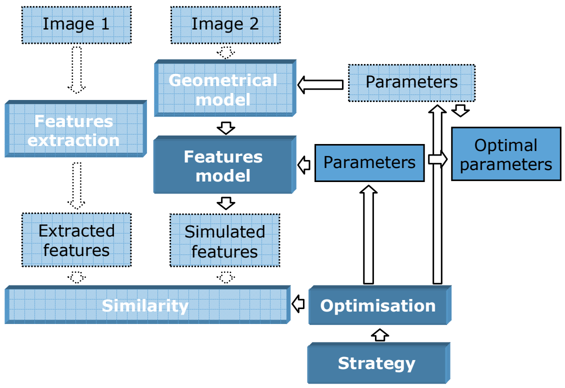



Figure 9. Scheme of a combine registration/scene interpretation technique.

## 5. CONCLUSIONS

After a short bibliography we have reminded the fact that automatic multisensor registration is nowadays almost operational for decametric resolution. Nevertheless we have underlined that it may be not an immediate issue for metric resolution.

Content difference can not be ignored in the coregistration procedure. We have demonstrated that a very coarse classification of ground and above-ground improves the coregistration accuracy.

Finally we propose to include the scene-interpretation process into the coregistration process rather than to perform it before or after this in order to benefit of the multisensor approach.

## Acknowledgements

This work has been partially supported by CNES funding.
The authors are grateful to ONERA and DGA for providing SAR data and to CNES for providing optical data.

**References:**

[1] Inglada, J. Giros, A., 2004. On the Possibility of Automatic Multisensor Image Registration. IEEE Transactions on Geoscience and Remote Sensing. Vol. 42, Issue 10, Oct. 2004 Page(s):2104 – 2120.

[2] Inglada, J.; Vadon, H, 2005. Fine registration of SPOT5 and Envisat/ASAR images and ortho-image production: a fully automatic approach. Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05. Volume 5, 25-29 July 2005 Page(s):3510 - 3512

[3] Hua-Mei Chen; Varshney, P.K.; Arora, M.K, 2003. Mutual information based image registration for remote sensing data. IEEE Transactions on Geoscience and Remote Sensing, Vol. 41, Issue 11, Part 1, Nov. 2003 Page(s):2445 - 2454

[4] Le Moigne, J.; Morisette, J.; Cole-Rhoades, A.; Netanyahu, N.S.; Eastman, R.; Stone, H, 2003. Earth science imagery registration. Proceedings IEEE International Geoscience and Remote Sensing Symposium, 2003. IGARSS '03.

[5] Wu, Y.; Maitre, H., 2000. Registration of a SPOT image and a SAR image using multiresolution representation of a coastline. Conference on Pattern Recognition, 1990. Proceedings., 10th International Vol. i, Date: 16-21 Jun 1990, Pages: 913 - 917 vol.1

[6] Chen, H., Varshney, P. K., 2000. A pyramid approach for multimodality image regisration based on mutual information. Proceedings of Fusion’ 2000, July 10-13, Paris, (International Society of Image Fusion), MoD3-9 – MoD3-15.

[7] Inglada, J, 2002. Similarity measures for multisensor remote sensing images. Proc. of Int. Geosciences And Remote Sensing Symposium (IGARSS02), Toronto, Canada.

[8] Roche, A., 1998. The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration. Technical report 3378, INRIA.

[9] Galland, F.; Tupin, F.; Nicolas, J.-M.; Roux, M., 2005. Registering of synthetic aperture radar and optical data. Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05. Vol. 5, 25-29 July 2005 Page(s):3513 - 3516

[10] Zhang, J. Wang, X., Chen, T., Zhang, Y., 2005. Change detection for the urban area based on multiple sensor information fusion. Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05. Vol. 1, 25-29 July 2005 Page(s):226 - 229

[11] Studholme, C, 1999. An Overlap Invariant entropy measure of 3D medical image alignment. Pattern Recognition, 32(1).

[12] A. Lorette, 1999. Texture Analysis through Markov Random Fields: Urban Areas Extractions. Proc. IEEE International Conference on Image Processing (ICIP), Kobe, Japan.

[13] Fischler, M. A., Bolles, R. C., 1981. Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography. Communication of the. ACM 24 (6), 381—395.