REMOTE SENSING AND IMAGE FUSION METHODS: A COMPARISON

Thierry Ranchin

Center for Energy and Processes, Observation, Modeling and Decision Team, Ecole des Mines de Paris
Rue Claude Daunesse, BP 207, 06904 Sophia Antipolis cedex, France.
e-mail: thierry.ranchin@ensmp.fr - Phone: +33 (0)4 93 95 74 53

ABSTRACT

By fusing two sets of images A and B, one with a high spatial resolution, the other with a low spatial resolution and different spectral bands, the ARSIS (from its French name: "amélioration de la résolution spatiale par injection de structures". It is usually translated by improvement of the spatial resolution by structures' injection) concept permits to synthesize the dataset B at the resolution of A with high quality. It tries to establish a relationship between the high frequencies in the multispectral set B and the set A and models it. The general problem for the synthesis is presented. General properties of the fused product are given. Then, the ARSIS concept is discussed. The general scheme for the implementation of a method belonging to this concept is presented. Three Inter-Band Structure Models (IBSM) are described. They are applied to Ikonos image as an illustration case. The fused products are assessed by means of a known protocol comprising a series of qualitative and quantitative parameters. Results and future tracks are discussed.

1. INTRODUCTION

This paper focuses on fusion of remotely sensed images. The general approach in the fusion of images is to create a new set of images, usually of reduced dimension, from the original sets of images. These sets may be commensurate or not and may originate from various modalities, e.g., panchromatic, microwave, hyperspectral, taken at different instants and with different times of integration and may have different space resolutions. Within a set, all images are geometrically aligned and have the same pixel size [1]. Here the term image comprises any information that is presented in a raster format, or gridded format in two dimensions. The grid cell is called pixel.

The objects of the fusion of images are various [1]. In environmental studies, a classical example is the classification process. Several images of commensurate or non-commensurate measurements and possibly of other information are inputs to a classifier. If the classification is of supervised type, a codebook exists that is input to the fusion process as an external knowledge. The result is an image of taxons and possibly another image of the related accuracy (or plausibility, or reliability etc.).

Many approaches in image fusion exist. This paper focuses on a few methods applying to remotely sensed images and imaging sensors for synthesis of images having the best spatial resolution available in the original sets of images.

One may distinguish three groups of methods [1]:

– projection and substitution methods: projection of original data sets into another space, substitution of one vector by the high resolution image and inverse projection into the original space, such as the IHS (Intensity, Hue, and Saturation) method [2],
– relative spectral contribution such as the Brovey transform [3] which can be applied to any set of image and the CNES P+XS method [4] dedicated to the SPOT case. It should be noted that the Brovey transform does not well represent this group because of its poor principles in construction [1]. Nevertheless, it is often used,
– methods relevant to the ARSIS concept: scale-by-scale description of the information content of both images and high-quality transformation of the multispectral content when increasing the spatial resolution [5, 6]. HPF (High Pass Filtering) method is part of this set of methods but does not usually provide quality transformation of the multispectral content.

This latter will be the focus of this paper. In part 2, the ARSIS concept is exposed. Then some elements on its implementation are discussed. An example is proposed with discussion on the quality of the synthesized products. Some future tracks are proposed.

2. THE ARSIS CONCEPT

Let denote the acquired images of lowest spatial resolution by B_l, and the images of highest spatial resolution by A_h. The subscripts l and h denote the spatial resolution of images B or A, i.e., low and high resolution, respectively.

The fusion methods aim at constructing synthetic images B^*_h, which are close to what would be observed by a similar multispectral sensor with a better spatial resolution. The methods should perform a high-quality transformation of the multispectral content of B_l, when increasing the spatial resolution from l to h. The general problem is relevant to the fusion of representations and is the creation of a new representation B* from the original representations A and B [1]:

$$B^* = f(A, B)$$

It may be seen as the inference of the information that is missing in the images B_{kl} for the construction of the synthesized images B^*_{kh}.

The ARSIS concept is based on the assumption that the missing information is linked to the high frequencies of the representations A and B. It searches a relationship between the high frequencies in B and A and models such a relationship. A method belonging to the ARSIS concept performs typically the following operations: (i) the extraction of a set of information from A, (ii) the inference of the information that is missing in the images B_{kl} using this extracted information and (iii) the construction of the synthesized images B^*_{kh}. The most recent methods perform a scale-by-scale description of the information content of both images and synthesis of the high-frequency information missing to transform the low spatial resolution images into high spatial resolution high spectral content images. [5] shows that many schemes can be accommodated within the ARSIS concept.

It is difficult to sketch the general scheme for the application of the ARSIS concept. In the methods HPF and others [7, 8, 9], the modeling of the missing information from the image A to the image B is performed on moving windows of these images themselves. It is possible to focus more on the modeling of the missing high frequencies, expressed by Fourier coefficients or wavelet coefficients or other appropriate spatial transform [1].

Figure 1 presents the general scheme that applies in the case of use of a multiscale model [5]. This case is used in the following for a better description of the ARSIS concept. Similar schemes can be used in other cases, where other tools or strategies are used. The following sections detail several implementations of the ARSIS concept following the scheme in Figure 1. Input to the fusion process are the images A at high spatial resolution (A_h, resolution n°1) and the spectral images B at low spatial resolution (B_{kl}, resolution n°2).

Three models appear in this scheme. The Multiscale Model (MSM) performs a hierarchical description of the information content relative to spatial structures in an image. An example of such a model for remotely sensed images is the combination of the wavelet transform and multiresolution analysis [10]. [5] provides practical details for the implementation of the algorithm of Mallat combined with a Daubechies wavelet. When applied to an image, the MSM provides one or more images of details that are the high frequencies and one image of approximation, i.e. the lower frequencies. As an example, assume an Ikonos image at 1 m resolution. The first iteration of the MSM gives one image of the structures comprised between, say, 1 and 2 m (details image) and one image of the structures larger than 2 m (approximation image). The spatial variability within an image can thus be modeled and the model can be inverted (MSM-1) to perform a synthesis of the high-frequency information to retrieve the original image at 1 m.

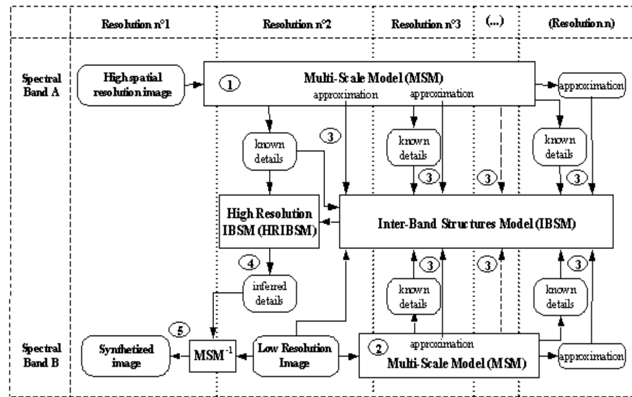

Fig. 1. General scheme for the application of the ARSIS concept using a multiscale model (MSM) and its inverse (MSM^{-1}).

The Inter-Band Structure Model (IBSM) deals with the transformation of spatial structures with changes in spectral bands. It models the relationships between the details or approximation observed in the representation A and those observed in the representation B. The IBSM may relate approximations and/or details for one or more resolutions and one or more spectral bands. As an example, the Model 2 relates the details observed at resolution n°3 in the image A and the image B_k by means of a linear relationship [5].

The High Resolution Inter-Band Structure Model (HRIBSM) performs the transformation of the IBSM with the change in resolution. This operation is not obvious. Many works have demonstrated the influence of the spatial resolution on the quantification of parameters extracted from satellite imagery. To our knowledge, no published fusion method paid particular attention to this point and the HRIBSM is often set identical to the IBSM. [11, 6] performed a multiscale synthesis of the parameters of their IBSM from resolution n°3 to resolution n°2.

3. IMPLEMENTATIONS OF THE ARSIS CONCEPT

In the general case, the operations are performed as follows. First, the MSM is used to compute the details and the approximations of image A (Step 1 in Figure 1). The same operation is applied to image B (Step 2). The analysis is performed for several resolutions, up to n in Figure 1 - that is (n-1) iterations for the analysis of the image A and (n-2) iterations for that of B_{kl}. These analyses provide one approximation image and several images of details for A and B. The known details at each resolution are used to adjust the parameters of the IBSM (Step 3). From this model is derived the HRIBSM (at resolution n°2 in this Figure), which converts the known details of image A into the inferred details of image B_k (inferred details, Step 4). Finally, MSM-1 from resolution n°2 to resolution n°1 performs the synthesis of the image B^*_{kh} (Step 5).

[5, 6] present examples of implementation of the different models. An implementation of the ARSIS concept is a coordinate choice of a MSM, an IBSM and a HRIBSM models. Several solutions are possible. A discussion on the appropriate choice of the models according to user needs, was conduced in [12].

The implementation of the proposed concept corresponds to the choice of a specific algorithm for each of the three models. [6] proposes a set of examples of the different implementations of the ARSIS concept.

In the following example, the selected MSM is the Undecimated Wavelet Transform (UWT) [13]

The first method is called UWT-RWM. The IBSM model used is the RWM (from the names of its authors Ranchin, Wald, Mangolini) model. This model is described in great details in [6]. It is a local adjustment of the statistics of the distributions of the details. This model is computed at resolution n° 4 (8 m). The HRIBSM model is identical to the IBSM model, and applied to resolutions n°2 and n° 3 (resp. 2 and 4 m).

The second method is called UWT-M1. In this case, the IBSM and HRIBSM models are the identity model.

The last one proposed in this paper is called UWT-M2. The IBSM model is model M2, a global adjustment of the distribution of the details modeled by the MSM, computed at resolution n°4 (8 m). The HRIBSM model is in this case identical to the IBSM model.

4. EXAMPLE AND RESULTS

The company GIM (http://www.gim.be) kindly provided an Ikonos dataset for test purposes. This set is composed of a panchromatic image Pan at the spatial resolution of 1 m and of four multispectral (MS) images at the spatial resolution of 4 m. The geographical area is the city of Hasselt in Belgium. The images were acquired April 28, 2000 at 10:39 UTC. The original data were delivered with a dynamic range of 11 bits (gray values). The application of the methods leads to the synthesis of MS images at the spatial resolution of 1 m with the same dynamic range in 11 bits.

The protocol of quantitative evaluation of the fused products defined by [14] was applied. This protocol is based on three properties:

- Any synthetic image B^*_h, once degraded to its original resolution l, should be as identical as possible to the original image B_l.
- Any synthetic image B^*_h should be as identical as possible to the image B_h that the corresponding sensor would observe with the highest resolution h.
- The multispectral set of synthetic images B^*_h should be as identical as possible to the multispectral set of images B_h that the corresponding sensor would observe with the highest resolution h.

All the computations were made on the full dynamic range of images (11 bits). The different quantitative results are not detailed in this paper. It was found that the three properties detailed [14], are respected with different levels of accuracy for the three methods.

As an illustration, Table 1 deals with the test of the second property and reports some statistics on the relative discrepancies between the original images B_{kl} and the images B^*_{kl}. The differences are computed on a pixel basis and one image of differences is obtained per spectral band *k*. From each image of differences, the mean value (bias), standard deviation and root mean square error (RMSE) are computed. In Table 1, these quantities are expressed in percent, relative to the mean radiance value of the original image B_{kl}. The ideal value for these parameters is 0. In addition, the difference between the variance of the original image B_{kl} and that of B^*_{kl} is computed. It is expressed in percent, relative to the variance of the original image. Ideally, this value should be zero. The correlation coefficient between the original image B_{kl} and B^*_{kl} is also computed. The ideal value is 1.

		Bias	Standard Dev.	RMSE	Diff. in Variance	Correlation Coeff.
Blue	UWT-RWM	0.01	3	3	23	0.92
	UWT-M1	0.02	8	8	-90	0.71
	UWT-M2	0.03	4	4	-20	0.87
Green	UWT-RWM	0.01	5	5	24	0.91
	UWT-M1	0.03	7	7	-22	0.84
	UWT-M2	0.04	5	5	-8	0.92
Red	UWT-RWM	0.01	8	8	23	0.92
	UWT-M1	0.03	10	10	-4	0.85
	UWT-M2	0.08	7	7	1	0.92
NIR	UWT-RWM	0.03	8	8	9	0.96
	UWT-M1	0.04	9	9	24	0.95
	UWT-M2	0.10	5	5	-1	0.98

Table 1. Some statistics of the relative differences and the relative difference in variance (all in percent) and the correlation coefficient between the original and synthesized images for the spectral bands.

For all methods, the bias is very small. This is in accordance with the first property. The standard deviations and RMSEs are small for all methods. The fused products UWT-RWM do not contain enough variance (*i.e.* fine structures). For the blue and green bands, the UWT-M1 and UWT-M2 add too many small structures (difference in variance is negative). This fact is confirmed by the low correlation coefficient in these cases. This observation meets the visual analysis performed on the 1 m products.

The UWT-RWM method is the method given the best achievable results whatever the spectral band (second property), as well as on the multispectral set of products (third property, results not shown here). The UWT-M2 method gives similar results. However, the addition of fine structures is usually more annoying for further processing than a lack of fine structures of similar intensity (difference in variance positive). Both can be considered as satisfactory. The UWT-M1 method is that giving the worst results of the four, the two latter being fairly unsatisfactory.

5. CONCLUSIONS AND FUTURE TRACKS

The ARSIS concept is an excellent framework for the development of accurate methods that can be tailored to specific users needs and can be assembled within a toolbox [15]. It is also an open framework with still many places for the development of different cases of applications and approaches for implementation. Different methods can be developed based on this concept, depending upon the multiscale description and synthesis model MSM, the IBSM relating the content of both representations A and B and the HRIBSM transforming the parameters of the IBSM when increasing the spatial resolution [1, 5, 6].

Future tracks of the development of the ARSIS concept are based on users needs and on the key problems encountered. Present works are oriented on the inclusion of the sensors parameters within the MSM models, on the development of metrics for construction of IBSM and HRIBSM models and on the refinement of the evaluation protocol in order to provide a normative approach for the evaluation of the synthesized products.

6. REFERENCES

[1] L. Wald, *Data Fusion. Definitions and Architectures - Fusion of Images of Different Spatial Resolutions*. Presses de l'Ecole, Ecole des Mines de Paris, Paris, France, 200 p. , 2002.

[2] W. J. Carper, T. M. Lillesand, and R. W. Kiefer, "The use of Intensity Hue Saturation transformations for merging SPOT panchromatic and multispectral image data." *Photogrammetric Engineering & Remote Sensing*, 56(4), 459-467, 1990.

[3] C. Pohl and J. L. van Genderen, "Multisensor image fusion in remote sensing: concepts, methods and applications." *International Journal of Remote Sensing*, 19(5), 823-854, 1998.

[4] Anonymous, *Guide des utilisateurs de données SPOT*, 3 tomes, Editeurs CNES et SPOT-Image, Toulouse, France, 1986.

[5] T. Ranchin, and L. Wald, "Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation." *Photogrammetric Engineering & Remote Sensing*, 66(1), 49-61, 2000.

[6] T. Ranchin, B. Aiazzi, L. Alparone, S. Baronti, and L. Wald, "Image fusion. The ARSIS concept and some successful implementation schemes." *ISPRS Journal of Photogrammetry & Remote Sensing*, 58(1-2), 4-18, 2003.

[7] J. G. Liu and J.M. Moore, "Pixel block intensity modulation: adding spatial detail to TM band 6 thermal imagery." *International Journal of Remote Sensing*, 19(13), 2477-2491, 1998.

[8] Y. Cornet, S. de Béthune, M. Binard, F. Muller, G. Legros and I. Nadasdi, "RS data fusion by local mean and variance matching algorithms: their respective efficiency in a complex urban context." *In* Proceedings of the IEEE/ISPRS joint Workshop on Remote Sensing and Data Fusion over Urban Areas, pp. 105-111, 2001.

[9] C. Diemer and J. Hill, "Local correlation approach for the fusion of remote sensing data with different spatial resolutions." *In* Proceedings of the Third Conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images", T. Ranchin and L. Wald Eds, published by SEE/URISCA, Nice, France, pp. 91-98, 2000.

[10] T. Ranchin, "Wavelets, remote sensing and environmental modelling." In Proceedings of the 15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics, Volume 6: Applications on Modelling and Simulation, pp. 27-34, 1997.

[11] T. Ranchin, L. Wald and M. Mangolini, "Efficient data fusion using wavelet transforms: the case of SPOT satellite images." *In* Proceedings of SPIE 1993 International Symposium on Optics, Imaging and Instrumentation. Wavelet Applications in Signal and Image Processing, Vol. 2034, pp. 171-178, 1994.

[12] T. Ranchin and L. Wald, "Different implementations of the ARSIS concept to fulfill users needs." *In* Proceedings of the 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas URBAN 2003, Berlin, Germany, 22-23 May 2003, Technical University of Berlin, pp. 299-304, 2003.

[13] P. Dutilleux, "An implementation of the "algorithme a trous" to compute the Wavelet Transform." In: Combes, J. M., Grossman, A., Tchamitchian, Ph. (Eds.), Wavelets: Time-Frequency Methods and Phase Space. Springer, Berlin, pp. 298-304, 1989.

[14] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images." *Photogrammetric Engineering & Remote Sensing*, vol. 63, no 6, pp. 691–699, 1997.

[15] C. Thomas, T. Ranchin, L. Wald, and J.-C. Laneri. "A modular platform for fusion of images." *In* Proceedings of the Fourth International Conference Physics in Image and Signal Processing PSIP'2005, January 31-February 2 2005, Toulouse, France, ISBN 2-912328-22-5, pp. 269-273, 2005.