Image Fusion - The ARSIS concept and some successful implementation schemes

Thierry Ranchin, Bruno Aiazzi, Luciano Alparone, Stefano Baronti, Lucien Wald

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

Image Fusion - The ARSIS concept and some successful implementation schemes

Thierry Ranchin[a], Bruno Aiazzi[b], Luciano Alparone[c], Stefano Baronti[b], Lucien Wald[a,*]

[a]Ecole des Mines de Paris, BP 207, 06904 Sophia Antipolis Cedex, France

[b]Institute of Applied Physics "Nello Carrara" IFAC-CNR, 50127 Firenze, Italy

[c]Department of Electronics and Telecommunications, University of Florence, 50139 Firenze, Italy

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

Abstract

This article aims at explaining the ARSIS concept. 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 concept permits to synthesise the dataset *B* at the resolution of *A* that is as close as possible to reality. It is based on the assumption that the missing information is linked to the high frequencies in the sets *A* and *B*. It searches a relationship between the high frequencies in the multispectral set *B* and the set *A* and models this relationship. The general problem for the synthesis is presented first. The 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. Then, this article intends to help practitioners and researchers to better understand this concept through practical details about

- Corresponding author: Tel.: +33-493957449; Fax: +33-493957535.

  E-mail address: lucien.wald@ensmp.fr (L. Wald).

implementations. Two Multi-Scale Models are described as well as two Inter-Band Structure Models. They are applied to an Ikonos image as an illustration case. The fused products are assessed by the means of a known protocol comprising a series of qualitative and quantitative tests. The products are found of satisfactory quality. This case illustrates the differences existing between the various models, their advantages and limits. Tracks for future improvements are discussed.

Keywords: multiresolution analysis; multiscale analysis; image pyramid; wavelet; data fusion; remote sensing

1. Introduction

Remote sensing images exhibit usually either high spectral resolution and low spatial resolution, or low spectral resolution (broadband) and high spatial resolution. The high spatial resolution is necessary for an accurate description of shapes, features and structures. The different objects are better identified if high spectral resolution images are used. Hence, there is a desire to combine the high spatial and the high spectral resolutions with the aim of obtaining the most complete and accurate description of the observed scene. Research has developed that aims at proposing algorithms for fusing both types of images, in order to synthesise images with the highest spectral and spatial resolutions available in the sets of images. Here, the vocabulary in data fusion recommended by the working group "Data Fusion" of the European Association of Remote Sensing Laboratories (EARSeL), a Regional (European) Member of the International Society for Photogrammetry and Remote Sensing (ISPRS), is adopted (Wald, 1998, 1999, 2002). A number of studies demonstrate the benefits of such fused products for the study of urban areas (André et al., 2002; Cornet et al., 2001; Couloigner et al., 1998a, b; Fanelli et al., 2001; Galaup and Pedron, 2002; Kishore Das et al., 2001; Terretaz, 1997; Ranchin and Wald, 1996a, b; Raptis et al., 1998; Vaiopoulos et al., 2001; Wald and Ranchin, 2001). The high-quality transformation of the spectral content of the multispectral images, when increasing the resolution, allows further processing such as the application of a classifier, automatic or not. For example, classification methods may be used as a first step to improve the extraction of streets (Ranchin and Wald, 1997).

Several well-known methods provide a better visual representation of the image (Carper et al., 1990; Mangolini, 1994; Pohl and van Genderen, 1998; Vrabel, 2000). They are very useful for photo-interpretation. This is particularly true when the number of spectral bands is much larger than the usual three bands for describing colours: red, green, blue. These "visual enhancement" may be performed to increase the utility of a set of images for visual analysis.

These methods have their limitations, especially with the new space-borne sensors and the demands from users for the reconstruction of high spatial resolution landscapes with objects having their natural colours. Here, in this context, natural colours mean the colours that are perceived by the human eye. Examples are the recent commercial space missions, Ikonos and QuickBird, that provide images with high spatial resolution images at respectively 1 and 0.7 m, and four multispectral images at a spatial resolution four times less (*i.e.* 4 and 2.8 m), taken in the blue, green, red and near infrared bands. The accurate synthesis of the multispectral character when increasing spatial resolution is very important to many applications, including those calling upon classification or the reproduction of the natural colours. Classification processes often use bases of spectra, which result from measurements or simulations by models or from the experience of image analysts. In the course of the classification, the observed spectra are compared to the known ones and a decision is taken according to their similarities. Accordingly, any error in the synthesis of the spectral signatures at the highest spatial resolution induces an error in the decision. These synthesised images should be as close as possible to reality and should simulate what would be observed by a sensor having the same spectral bands but the highest spatial resolution.

This article is concerned only with those methods which claim to provide synthetic images close to reality when enhancing the spatial resolution by fusion of two types of images, and not those which only provide a better visual representation of the dataset. Several synthesis methods by fusion have been published or are available in commercial software packages. Ranchin and Wald (2000a) distinguish three groups of methods: the projection and substitution methods, the relative spectral contribution methods and those relevant to the ARSIS (from its French acronym: “Amélioration de la Résolution Spatiale par Injection de Structures”, which means “spatial resolution enhancement by injection of structures”) concept. Evidently, there are some hybrid methods belonging to more than one group (e.g., Nuñez et al., 1999). Several investigations demonstrated that the best presently achievable results are attained by the methods belonging to the ARSIS concept (Aiazzi et al., 1999;

Fanelli et al., 2001; Ranchin and Wald, 2000a, b; Raptis et al., 1998; Terretaz, 1997; Wald, 2002; Yang et al., 2000).

This article aims at explaining the concept of injecting high-frequencies in multispectral bands (using the ARSIS concept) and to demonstrate its advantages. First, the ARSIS concept and its properties are presented. Then, several successful schemes for implementation of the concept are provided and illustrated through an example of an Ikonos image of the city of Hasselt, Belgium. It discusses future tracks for improvement.

2. Problem statement

Let us 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. $B^{interp}{}_h$ denotes the result of the interpolation (resampling) of B_l from resolution l to h. Within each set, the images are geometrically aligned and have the same pixel sizes. Within the set of images B, B_{kl} denotes the image acquired in the spectral band k. The fusion methods aim at constructing synthetic images B^*_h, which are close to reality. 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 the creation of a new set of images B^* from the original sets of images

$$B^* = f(A, B) \qquad (1)$$

In addition, these synthetic images B^* must respect the three following properties (Wald et al., 1997).

2.1 First property

Any synthetic image B^*_h once degraded to its original resolution l, should be as identical as possible to the original image B_l, that is

$$D_l(B_{kl}, B^*_{kl}) < \varepsilon l_k \qquad (2)$$

where D_1 is the distance between B_{kl} and B^*_{kl}. Approximation induced by the resampling of B^*_{kh} into B^*_{kl} should be taken into account: the limit $\varepsilon 1_k$ is determined by the requested degree of accuracy. $\varepsilon 1_k$ should be small for all k; this ensures the similarity between the sets B_l and B^*_l. An example of D_1 is the root of the mean of the squared differences $(B_{kl} - B^*_{kl})$ on a pixel basis. A typical value for $\varepsilon 1_k$ is 0.05 times the mean value of B_{kl}. As for the other properties, other distances may be used in order to enhance specific properties in the images, e.g., structures, or to describe *local* inaccuracies, i.e. occurring at the pixel scale, that can impair the fused product.

2.2 Second property

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 spatial resolution h, if existent:

$$D_2(B_{kh}, B^*_{kh}) < \varepsilon 2_k \qquad (3)$$

where D_2 is the distance between B_{kh} and B^*_{kh} for k. As for $\varepsilon 1_k$, the limit $\varepsilon 2_k$ is determined by the requested degree of accuracy. The smaller $\varepsilon 2_k$, the greater the similarity between the sets B_h and B^*_h. If as previously, D_2 is the root of the mean of the squared differences, a typical value for $\varepsilon 2_k$ is 0.05 times the mean value of B_{kh}.

2.3 Third property

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 spatial resolution h, if existent:

$$D_3(B_h, B^*_h) < \varepsilon 3 \qquad (4)$$

where D_3 is the distance between the sets B_h and B^*_h. $\varepsilon 3$ is the limit set by the requested degree of accuracy. An example of D_3 is the ERGAS quantity, discussed later.

The ability to geometrically superimpose images is important, especially since the ARSIS concept is dealing with the addition / combination of high frequencies. The images B_l and A_l

should be geometrically aligned, once all images (here A) are degraded to the lowest available spatial resolution. Some systems provide images of different spatial resolutions that are already co-registered, such as Landsat images. Otherwise, this can be done by means of standard methods available in public or commercial software packages for image processing. Some providers of images arrange for their products to be co-registered. The images of lowest resolution B_l are projected into the geometry of A_l. During the process, a resampling of the multispectral images B is made. A few authors have assessed the influences of respectively the quality of the co-registration and the resampling operator on the final results (Blanc et al., 1998; Wald et al., 1997). The discrepancies between the results relative to the mean radiance (or grey value) of the actual images are a few per cent; these influences can be kept very small provided the co-registration is accurate enough and the operator is appropriate enough. In most cases, a bicubic interpolator offers a good compromise between the accuracy of the result and the required computer time. In the following, for the sake of the simplicity, the term "image of lowest resolution" B_l will denote the projected resampled image of lowest resolution.

3. The ARSIS concept

The general problem may be seen as the inference of the information that is missing in the images B_{kl} and the construction of the synthesised images B^*_{kh}. The ARSIS concept is based on the assumption that the missing information is linked to the high frequencies of the sets A and B. It searches a relationship between the high frequencies in the multispectral set B and the set 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 the set 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 synthesised 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. Ranchin and Wald (2000a) showed that many schemes can be accommodated within the ARSIS concept. Among them are the High-Pass Filtering (HPF) method (Chavez et al., 1991), the method by Aiazzi et al. (1999) and three models presented in Ranchin and Wald (2000a), making use of wavelet

transform: Model 1, Model 2 and RWM, the latter being named after the initials of its authors (Ranchin, Wald, Mangolini, see Ranchin et al. (1994)).

The images of the sets *A* and *B* do not need to be commensurate. Some studies have been published where images acquired in thermal infrared bands have been synthesised with a better spatial resolution with a satisfactory quality by the means of images acquired in the visible range (Kishore Das et al., 2001; Liu and Moore, 1998; Nishii et al., 1996; Wald and Baleynaud, 1999).

It is difficult to sketch the general scheme for the application of the ARSIS concept. In the methods HPF and by Cornet et al. (2001), Diemer and Hill (2000), Liu and Moore (1998), Pradines (1986) or Price (1999), the modelling 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 modelling of the missing high frequencies, expressed by Fourier coefficients or wavelet coefficients or other appropriate spatial transform.

Fig. 1 presents the general scheme that applies in the case of use of a multiscale model. 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 Fig. 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 (Ranchin, 1997). Ranchin and Wald (2000a) provide details for the implementation of the algorithm of Mallat (1989) combined with a Daubechies wavelet. When applied to an image, the MSM provides one or more images of details, that is the high frequencies, and one image of approximation, that is 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 modelled and the model can be inverted (MSM^{-1}) to perform a synthesis of the high-frequency information.

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 image A and those observed in the image 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 described by Ranchin and Wald (2000a) relates the details observed at resolution n°3 in the image A and the image B_k by means of a linear relationship. Mangolini (1994) reported attempts to relate details observed at two successive resolutions.

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 (Lillesand and Kiefer, 1994; Woodcock and Strahler, 1987). To our knowledge, no published fusion method paid particular attention to this point and the HRIBSM is often set identical to the IBSM. Ranchin et al. (1994) performed a multiscale synthesis of the parameters of their IBSM from resolution n°3 to resolution n°2.

The operations are performed as follows. First, the MSM is used to compute the details and the approximations of image A (Step 1 in Fig. 1). The same operation is applied to image B (Step 2). The analysis is performed for several resolutions, up to n in Fig. 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).

4. Multi-Scale Models: generalised Laplacian pyramid and “à trous” wavelet transform

MSM performs a hierarchical description, modelling and synthesis of the information content relative to spatial structures in an image. Fig. 2 is a description of pyramidal algorithms and more generally of multiscale models (Mallat, 1989). The basis of the pyramid is the original image. Each level of the pyramid is an approximation of the original image

computed from the original one. When climbing the pyramid (the analysis), the successive approximations have coarser and coarser spatial resolution. The computation of the approximations is done using a base of functions, called the scale functions. The basis of the pyramid is the original image provided by the sensor. In this scheme, the use of the Laplacian pyramid or wavelet transform allows the description of the differences existing between two successive approximations of the same image (*i.e.* two successive levels of the pyramid). These differences are called details. If the process of the multiresolution analysis is inverted, the original image can be exactly reconstructed, from one approximation and from the different details describing the differences in signal between this approximation and the original image: this is called synthesis.

Nuñez et al. (1999) describe the implementation of the “à trous” wavelet transform. This method was first proposed by Kronland-Martinet et al. (1987) for music synthesis. The term “à trous” ("with holes") was introduced by Dutilleux (1989) and relates to the fact that the even coefficients of the used scaling function (and the associated wavelet function), except the central one, are zero. A theoretical analysis of the “à trous” method is given in Shensa (1992). It is a nonorthogonal, shift-invariant, symmetric, dyadic, undecimated, discrete wavelet transform. Practically, it leads to a band-pass stack of images with same dimensions (no decimation, thus no pyramid), with reduction of resolution by factor 2 from level to level, using a Gaussian-like low-pass filter. 2^i - 1 zeros are inserted between each pair of the filter coefficients, when filtering level i (i = 0 for the original image). The wavelet planes are given by the difference of two consecutive levels of the stack. The filter coefficients (and the wavelet function) are computed based on the selection of the scaling function. Usually, a cubic B-spline scaling function is selected resulting in a 5 x 5 low-pass filter (Nuñez et al., 1999). From the above, it is obvious that “à trous” is an Undecimated Wavelet Transform (UWT). The implementation of the Generalised Laplacian pyramid (GLP) is detailed hereafter. Common feature of both is that they are *redundant* or *oversampled* multiresolution transforms, which provide significant benefits for image fusion, as demonstrated by Aiazzi et al. (2002).

The Laplacian pyramid (LP) is derived from the Gaussian pyramid (GP) which is a sequence of multiresolution approximations obtained through a recursive *reduction* of the image dataset. Reduction by 2 ($\text{reduce}_2(.)$) is separable low-pass filtering followed by

decimation by 2 along rows and columns. The 2D low-pass reduction filter is generally zero-phase (Aiazzi et al., 2002). If $G_0(m,n)$ is a grey-scale image, the GP is defined as

$$G_k(m,n) = \text{reduce}_2[G_{k-1}](m,n) \quad (5)$$

where k identifies the level of the pyramid, K being the top level (approximation).

From the GP, the *enhanced* LP (ELP) (Aiazzi et al., 1997) is defined as

$$L_k(m,n) = G_k(m,n) - \text{expand}_2[G_{k+1}](m,n) \quad (6)$$

in which $\text{expand}_2[G_{k+1}]$ denotes the (k+1)st GP level expanded by 2 to match the size of the underlying kth level. The 2D low-pass filter for expansion is still separable and zero-phase and must cut-off at one half of the signal bandwidth to exactly reject the spectral images introduced when samples are zero-interleaved. The baseband approximation is added to the band-pass ELP, i.e. $L_K(m,n) = G_K(m,n)$, to yield a complete image description comprising both approximation and details.

When the scale ratio is not a power of 2, but any integer or fractional number, the operators $\text{reduce}_2(.)$ and $\text{expand}_2(.)$ may be generalized to deal with fractional reduction and expansion (Aiazzi et al., 1999). The outcome pyramid will be denoted as GLP, irrespective of scale ratio.

5. IBSM models

5.1 The model of Aiazzi, Alparone, Baronti and Pippi (AABP)

The model IBSM deals with the transformation of spatial structures with changes in spectral bands. It models the relationships between the details or context observed in the image A and those observed in the image B. The AABP model is noted from its authors' initials (Aiazzi et al., 2001).

Let C_B^l be the details obtained from the application of the Multi-Scale Model on the image B in the spectral band k under concern at resolution l. That is C_B^l represent the transitions (details) for sizes comprised between l and $2l$. Same notations for C_A^l for the image A. The

AABP model establishes a local relationship between the details C_B^l and the details C_A^l based on a local relationship between the images (approximations) A_l and B_{kl}.

We assume that the following relationships hold, where C_B^h are the high-frequency details necessary for the construction of the high resolution multispectral image B_{kh}:

$$C_B^l = a^l C_A^l \text{ (IBSM)} \qquad \text{and} \qquad C_B^h = a^l C_A^h \text{ (HRIBSM)} \tag{7}$$

where a^l is computed on a sliding window of 7 x 7 (SPOT 1 to 4 case) or 9 × 9 (Ikonos case) at the resolution h and depends upon the spectral band k. In the following, σ_A and σ_B are respectively the standard deviations of A_l and B_{kl} and ρ is the linear correlation coefficient of Pearson (CC) for the n x n window. Let also θ be a constant threshold ranging in 0.3-0.6 depending on the global cross-correlation between A_l and B_{kl}, the lower the correlation, the higher the threshold. To avoid numerical instabilities on homogeneous areas of A, a^l is clipped above 3 and thus given by:

$$a^l = min[\sigma_B / (1+\sigma_A), 3] \quad \text{if } \rho \geq \theta \tag{8}$$

and

$$a^l = 0.0 \qquad \text{if } \rho < \theta \tag{9}$$

5.2 RWM model

The notations are the same as before. The RWM model establishes a local relationship between the details C_B^l and the details C_A^l. Compared to the previous AABP model, this model is also context-driven but in the space of the details and not of approximations.

We assume that the following linear relationships hold, where C_B^h are the details necessary for the construction of the high resolution multispectral image *Bkh*:

$$C_B^l = a^l C_A^l + b^l \text{ (IBSM)} \tag{10}$$

and

$$C_B^h = a^l\, C_A^h + b^l \text{ (HRIBSM)} \qquad (11)$$

where the gain a^l and offset b^l are the parameters of the first axis of inertia computed on a sliding window of size n lines and n columns at the resolution l and are function of the spectral band. According to the experience gained in the SPOT-1 case, n is set to $(7(2\,l) / h + 1)$ pixels of size h.

In the following, m_{CB} and m_{CA} are respectively the mean of C_B^l and C_A^l, σ_{CB} and σ_{CA} respectively the standard deviations of C_B^l and C_A^l, cov is the covariance between C_B^l and C_A^l and ρ is the linear correlation coefficient of Pearson for the window nxn. In order to avoid noise, only the most energetic transitions are considered. The standard deviations, covariance and correlation coefficient are computed for the coefficients whose absolute values are greater than the absolute value of the mean. To avoid numerical problems, ρ is only computed if $\sigma_{CB}\,\sigma_{CA}$ is greater than 10^{-5} and $|cov|$ greater than 10^{-2}. Let us denote *sign* the sign of the product $C_B^l\, C_A^l$ for the central pixel.

Two positive thresholds are defined, s_1 and s_2, with s_1 greater than s_2. s_1 and s_2 are set respectively to 0.7 and 0.01. There are three cases. Let us define

$$\alpha 0 = sign\ \sigma_{CB} / \sigma_{CA} \text{ and } \beta 0 = m_{CB} - \alpha 0\, m_{CA} \qquad (12)$$

First case: $|\rho| < s_2$

if $\sigma_{CA} < 10^{-5}$ or $\rho = 0$, $\alpha 0 = 0$

$$\alpha = \alpha 0 \text{ and } \beta = \beta 0 \qquad (13)$$

Second case: $|\rho| \geq s_1$

$$\alpha = \frac{(\sigma_{CB}^2 - \sigma_{CA}^2) + \sqrt{(\sigma_{CB}^2 - \sigma_{CA}^2)^2 + 4\mathrm{cov}^2}}{2\mathrm{cov}} \text{ and } \beta = m_{CB} - \alpha\, m_{CA} \qquad (14)$$

If $|cov| \leq 10^{-2}$ $\alpha = \alpha 0$ and $\beta = \beta 0$ (15)

Third case: $s_1 > |\rho| \geq s_2$

$$\alpha = \alpha 0 + \left(\frac{(\sigma_{CB}^2 - \sigma_{CA}^2) + \sqrt{(\sigma_{CB}^2 - \sigma_{CA}^2)^2 + 4\mathrm{cov}^2}}{2\mathrm{cov}} - \alpha_0 \right) \left(\frac{|\rho| - s_2}{s_1 - s_2} \right)$$

and $\beta = m_{CB} - \alpha\, m_{CA}$ *(16)*

If $|cov| \leq 10^{-2}$ $\alpha = \alpha 0$ and $\beta = \beta 0$ *(17)*

Taking into account the calibration coefficients, d_B and d_A, of respectively the multispectral and panchromatic sensors converting image grey values into radiances, we obtain finally

$a^l = \alpha\ d_B / d_A$ and $b^l = \beta\ d_B$ *(17)*

6. A case study: an Ikonos dataset of Hasselt, Belgium

The company GIM kindly provided an Ikonos dataset for test purposes. This dataset comprises a panchromatic image PAN with a resolution of 1 m and four spectral images with a resolution of 4 m. The spectral bands are given in Table 1. The grey values are coded in 11 bits. The images were given by Space Imaging as being geocoded and superimposable but discrepancies of up to 20 m were found between PAN and multispectral images. Accordingly, they were again co-registered by the means of an automatic method based on multi-resolution analysis and local deformation models (Blanc and Wald, 1998).

The geographical area is the city of Hasselt in Belgium. The images were taken simultaneously on 28 April 2000, at 10:39 UT. Table 1 reports the mean value and standard deviation of each band. It also provides the correlation coefficient between each band and the panchromatic image resampled at 4 m.

Fig. 3a shows a region of the PAN image at 1 m. On the left, there is a river, crossed by two bridges, with small boats and several barges. Along the river bank is a main street. Several cars are visible. This area is mainly an industrial district with large buildings, surrounded by numerous trees. On the top right is a stadium. Its lawn is partly degraded.

The two models MSM and the two models IBSM are combined to offer three different methods that are implementations of the ARSIS concept. These methods are noted: GLP-AABP, UWT-AABP and UWT-RWM. The combination GLP-RWM cannot be performed without extensive changes in the RWM model. These three implementations were run on the Ikonos dataset.

7. Results and assessment

The results of the different methods are assessed by the means of the protocol proposed by a joint working group of EARSeL and of the Société des Electriciens et Electroniciens, the French branch of the IEEE (Wald et al., 1997).

This protocol permits to alleviate the need for a reference image if not available and offers a complete check of the three properties of fused products. It comprises a series of qualitative and quantitative tests (Wald, 2002). In this particular case, the protocol is as follows:

- the fusion method is applied to the original sets of images A_h and B_{kl}. It results into a new set of synthesised images B^*_{kh} at resolution of 1 m.
- testing the first property: *any synthetic image B^*_{kh}, once degraded to its original resolution l, should be as identical as possible to the original image B_{kh}.* To achieve this, the synthetic image B^*_{kh} is spatially degraded to an approximate solution $(B^*_{kh})_l$ of B_{kl}. If the first property is true, then, $(B^*_{kh})_l$ is very close to B_{kl}. The difference between both images is computed on a per-pixel basis. The fused products together with the difference images are visually compared to the original images B_{kl} in order to detect trends of error, if any, possibly related to the objects in the scene. Then, some statistical quantities are computed to quantitatively express the similarities and discrepancies between both images.
- a change of scale is performed for the second and third properties. Two sets of images A_l and B_{kv} are created from the original sets of images A_h and B_{kl}. The image A_h is degraded to the low resolution l (A_l, i.e. an Ikonos panchromatic image at 4 m) and the images B_{kl} to a lower resolution v (B_{kv}, i.e. an Ikonos multispectral image at 16 m), where $v=l(l/h)$,

- the fusion method is applied to these two sets of images, resulting into a set of synthesised images B^*_{kl} at resolution l. The original images B_{kl} (the Ikonos original multispectral image) serve now as references.
- the second and third properties are tested with the synthetic images B^*_{kl}. The quality observed for the fused products B^*_{kl} is assumed to be close to the quality that would be observed for the fused products B^*_{kh} if a reference at resolution h was present. This point has been largely discussed by Wald et al. (1997). A comparison is performed between B_{kl} and B^*_{kl} by the means of visual analysis and analysis of the similarities and discrepancies. For the third property, the emphasis is put on the spectral similarities.

The images synthesised at 1 m by the various methods are presented in Figs. 3b and c for the same geographical area as Fig. 3a and for the near-infrared band (NIR). To save space, the image produced by the method UWT-AABP is not shown; its visual appearance is very similar to those of the images GLP-AABP and UWT-RWM. In theory, these images cannot be compared to references since the latter are not available. However, given the importance of the NIR signal in the PAN band, a cautious comparison may be performed with the PAN image in Fig. 3a. The images GLP and UWT appear too smooth. The objects are not enough contrasted. A kind of halo seems to surround many objects in the UWT image; see in particular the boats and the stadium. That may translate an insufficient modelling of the details by the UWT. However, the UWT image offers better visualisation of other structures than the GLP image; see e.g. the elongated structure, parallel to the river, upper left of the middle of the picture.

The visual analysis is complemented by quantitative assessment, where some statistical quantities are computed to express the similarities and discrepancies between the fused images and the reference images (see discussion above). For the first and second properties, several quantities are computed (see e.g. Wald et al., 1997). For the sake of the simplicity, only reported here are the bias, the standard deviation of the differences and the root mean square error (RMSE), as well as the difference between the actual variance and the estimate and the correlation coefficient between the actual image B and the estimate B^*.

In testing the first property, an important point is the way the synthetic image B^*_{kh} is degraded to $(B^*_{kh})_l$ since the results depend on the filtering operator used. Wald et al. (1997) showed that the discrepancies between the results relative to the mean grey value of the

original image are on the order of a very few per cent. In conclusion, there is an influence of the filtering operator upon the results, but it can be kept very small provided the operator is appropriate enough. For the sake of concision, the verification of this property may be summarised in a few sentences instead of a table. The methods discussed here are by essence built to satisfy this first property. The discrepancies between the original images and the images $(B^*_{kh})_l$ computed on a pixel basis are very small, with reservations regarding the degradation process. For example, the standard deviation of the discrepancies in the NIR band relative to the mean value of the original image is less than 1 %.

Testing the second property reveals the properties of the methods (Fig. 4). Table 2 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. From each image of differences, the mean value (bias), standard deviation and root mean square error (RMSE) are computed. In Table 2, these quantities are expressed in percent, relative to the mean radiance value of the original image B_{kl}. The ideal values 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.

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 do not contain enough variance. This observation meets the visual analysis performed on the 1-m products. The three products are satisfactory. The best results are attained by the GLP-AABP products. The UWT-RWM comes second and then, the UWT-AABP. Taking into account the analysis of the images at 1 m, these observations are interpreted as follows. The superiority of the GLP-AABP originates from a better MSM. The superiority of the UWT-RWM originates from a more precise IBSM. In the case presented here, a precise MSM is better than a precise IBSM.

The first and second properties deal with the spectral bands individually. The third property may be tested by visually comparing colour composites made from the sets B^*_{kl} and B_{kl}. A quantitative assessment can also made to quantify the performance of the methods in synthesising the spectral signatures during the change in spatial resolution. These indications are of particular importance in case of classification calling upon spectral libraries or in the

case of production of images in true color. In other cases, since the spectra are taken out of the multispectral images themselves, this table does not mean at all that classification of fused products will lead to bad results. On the contrary, fused products based upon the ARSIS concept usually lead to enhanced mapping of classes (Couloigner et al., 1998b; González de Audicana and Seco, 2003; Raptis et al., 1998; Yang et al., 2000). A number of criteria were proposed by Wald (2002) and Wald et al. (1997) for the visual and quantitative assessment of the spectral fidelity of the fused products at 4 m compared to the original images. They deal with the number of dominant *n*-tuples, their synthesis and their frequencies.

Wald (2002) proposed an error that offers a global picture of the quality of a fused product. This error is called ERGAS, after its name in French "*erreur relative globale adimensionnelle de synthèse*" (dimensionless global relative error of synthesis). It is given by:

$$ERGAS = 100 \frac{h}{l} \sqrt{\frac{1}{N}\sum_{k=1}^{N}\frac{(RMSE(B_k))^2}{M_k^2}} \qquad (19)$$

This error is reported in Table 3 for the various cases. From several published and unpublished experiments, Wald (2002) reports that an error ERGAS larger than 3 corresponds to fused products of low quality, while an ERGAS less than 3 denotes a product of satisfactory quality or better. The three methods exhibit an ERGAS smaller than 3.

8. Conclusion

The ARSIS concept is a general framework for the improvement of the spatial resolution of multispectral images. It is now well understood and is now employed in applications such as urban mapping (see references in introduction), air quality in cities (Wald and Baleynaud, 1999), or agriculture (González de Audicana and Seco, 2003). It is a good and open framework with room for the further development of different applications and implementation approaches. The application of this concept leads to the construction of high spatial resolution multispectral images that are close to the images that the corresponding sensor would observe with the highest resolution. Different methods can be developed based on this concept, depending upon the multiscale description and synthesis model MSM, the

model ISBM relating the content of both datasets and the model HRIBSM transforming the parameters of the model IBSM when increasing the spatial resolution.

Practical details have been given for the implementation of several methods belonging to the ARSIS concept. These implementations are illustrated by a particular case. Several aspects were assessed: visual and performances in synthesising individual spectral images. These aspects are the most important with respect to the subsequent application of classification techniques on the synthesised multimodalities ensemble. It was found that all the methods were producing satisfactory results. This article is dealing with only one case and it is premature to draw firm conclusions regarding the compared properties of the methods.

Regarding the characteristics of Ikonos data, we noticed that the correlation coefficients between P and the MS bands (Table 1) are extremely low; they are less than 0.5 except for the NIR (0.875). Given the bandwidth characteristics of P and MS channels, this is a clue that the P band has been radiometrically processed in a nonlinear way, same as a photographic image, e.g. by applying a gamma to the grey values. Both the RWM model and the AABP model rely on local covariance measurements between P and each of the MS bands; therefore, they are penalized in performance by the nonlinear mapping of the P levels. A solution would be to try to estimate the non-linear transformation of P, invert it and use the remapped P to perform fusion with MS. This will be a possible object for future experiments.

Other aspects in assessment may have been considered, such as spatial gradients, shapes and structures, both in each spectral band and in the multispectral set. Such aspects and the corresponding criteria are of high importance in several applications such as the automatic recognition of objects, features, networks and so on. They have not been considered here. Hints about the performances of each method vis-à-vis these more specific aspects may be drawn from the present discussion. These aspects as well as others may have importance in the selection of a method in a given case.

It was observed that the GLP modelling provides better results than the UWT modelling. The superiority is partly hampered by an increase of accuracy in the IBSM model (UWT-RWM vs. UWT-AABP). Though the general scheme of the ARSIS concept differentiates the models MSM, IBSM and HRIBSM, this example demonstrates that these models are related. They should be designed in respective agreement for better results. The two presented IBSM models (AABP, RWM) are modelling each spectral band separately (as well as the HRIBSM

models). The example tends to show that better results would be attained if multispectral properties are taken into account in the IBSM and HRIBSM.

There are several ways of improvement. One is the choice of the multiscale analysis underlying the modelling and injection of spatial details. Several tools exist for the multiscale analysis and for the modelling of the high frequencies in the time-frequency domain. They have different properties and some may be more adapted than others, resulting in a better quality of the synthesised images. However, the rationale of spatial frequencies spectrum substitution from an image to another may help devise new analysis tools (Argenti and Alparone, 2000) that are suitable for specific applications, e.g. ground scales of the data whose ratio is a fractional number (Aiazzi et al., 2000).

The second way is expected to bring definite improvements. The modelling of the inter-modality behavior of the small-size structures (high frequencies) is central in the ARSIS concept (IBSM model). The models presently available are rather straightforward. Though they already produce satisfactory results, efforts should be made to improve them and finally provide better synthesised images. They are mostly based upon statistical adjustment of some properties representing the signal dynamics. Physical laws should be taken into account in these models. Efforts should also be made on the HRIBSM model, for which very few studies were performed, thus its behaviour is poorly known. It is believed that the improvement of the IBSM model will lead to improvements in the HRIBSM model.

Acknowledgements

The Ikonos image of Hasselt was kindly provided by the company GIM. This article partly synthesises several works performed under the auspices of the companies Aérospatiale, Alcatel Space Industries and SPOT-Image, the French Ministry of Defence and the French Space Agency CNES and the European Space Agency.

References

Aiazzi, B., Alparone, L., Baronti, S., Lotti, F., 1997. Lossless image compression by quantization feedback in a content-driven enhanced Laplacian pyramid. *IEEE Transactions on Image Processing* 6(6), 831-843.

Aiazzi, B., Alparone, L., Argenti, F., Baronti, S., 1999. Wavelet and pyramid techniques for multisensor data fusion: a performance comparison varying with scale ratios. In: Serpico, S. B. (Ed.), Image and Signal Processing for Remote Sensing V, SPIE Vol. 3871, pp. 251-262.

Aiazzi, B., Alparone, L., Argenti, F., Baronti, S., Pippi, I., 2000. Multisensor image fusion by frequency spectrum substitution: subband and multirate approaches for a *3:5* scale ratio case. Proceedings IEEE International Geoscience and Remote Sensing Symposium, pp. 2629-2631.

Aiazzi, B., Alparone, L., Baronti, S., Pippi, I., 2001. Quality assessment of decision-driven pyramid-based fusion of high resolution multispectral with panchromatic image data. In: Proceedings of the IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Rome, Italy, November 8-9th, pp. 337-341.

Aiazzi, B., Alparone, L., Baronti, S., Garzelli, A., 2002. Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Transactions on Geoscience and Remote Sensing* 40(10), 2300-2312.

André, G., Chiroiu, L., Guillande, R., Galaup, M., 2002. Evaluation et cartographie de dommages par imagerie satellitaire SPOT 5 : simulation sur la ville de Bhuj, séisme de Gujarat, Inde (26 janvier 2001). *Bulletin SFPT*, No 164/165**,** 174-183.

Argenti, F., Alparone, L., 2000. Filterbanks design for multisensor data fusion. *IEEE Signal Processing Letters* 7(5), 100-103.

Blanc, Ph., Wald, L., 1998. Validation protocol applied to an automatic co-registration method based on multi-resolution analysis and local deformation models. *International Archives of Photogrammetry and Remote Sensing*, vol. 32, Part 2, pp. 11-19.

Blanc, Ph., Wald, L., Ranchin, T., 1998. Importance and effect of co-registration quality in an example of " pixel to pixel " fusion process. In: Proceedings of the 2nd International Conference " Fusion of Earth Data : merging point measurements, raster maps and remotely sensed images ", Sophia Antipolis, France, 28-30 January. In: Ranchin, T., Wald, L. (Eds.), published by SEE/URISCA, Nice, France, pp. 67-74.

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

Chavez, P. S. Jr., Sides, S. C., Anderson, J. A., 1991. Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic. *Photogrammetric Engineering & Remote Sensing* 57(3), 265-303.

Cornet, Y., de Béthune, S., Binard, M., Muller, F., Legros, G., Nadasdi, I., 2001. 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, Rome, Italy, November, 8-9th, pp. 105-111.

Couloigner, I., Ranchin, T., Wald, L., 1998a. Benefit of data fusion to urban roads mapping. In: Proceedings of the second conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images", Sophia Antipolis, France, January 28-30, Ranchin, T., Wald, L. (Eds.), published by SEE/URISCA, Nice, France, pp. 183-190.

Couloigner, I., Ranchin, T., Valtonen, V. P., Wald, L., 1998b. Benefit of the future SPOT 5 and of data fusion to urban mapping. *International Journal of Remote Sensing* 19(8), 1519-1532.

Diemer, C., Hill, J., 2000. Local correlation approach for the fusion of remote sensing data with different spatial resolutions. In: Ranchin, T., Wald, L. (Eds.), Proceedings of the third Conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images", Sophia Antipolis, France, January 26-28. SEE/URISCA, Nice, France, pp. 91-98.

Dutilleux, P., 1989. 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.

Fanelli, A., Leo, A., Ferri, M., 2001. Remote sensing images data fusion: a wavelet transform approach for urban analysis. In: Proceedings of the IEEE/ISPRS joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Rome, Italy, November, 8-9th, pp. 112-116.

Galaup, M., Pedron, C., 2002. Mise à jour de plans d'informations de la base de données urbaines de la ville de Toulouse. *Bulletin SFPT* 164/165, 106-117.

González de Audicana, M., Seco, A., 2003. Fusion of multispectral and panchromatic images using wavelet transform. Evaluation of crop classification accuracy. In: Benes, T., (Ed.), Proceedings of 22nd EARSeL Annual Symposium "Geoinformation for European-wide

integration", 4-6 June 2002, Prague, Czech Republic. Millpress, Rotterdam, Netherlands, pp. 265-272.

Kishore Das, D., Gopal Rao, K., Prakash, A., 2001. Improvement of effective spatial resolution of thermal infrared data for urban landuse classification. In: Proceedings of the IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Rome, Italy, November, 8-9th, pp. 332-336.

Kronland-Martinet, R., Morlet, J., Grossman, A., 1987. Analysis of sound patterns through wavelet transforms. International Journal of Pattern Recognition and Artificial Intelligence 1(2), 273-301.

Lillesand, T. M., Kiefer, R. W., 1994. *Remote Sensing and Image Interpretation*. Third edition, John Wiley & Sons, New York, 750 p.

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

Mallat, S. G., 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(7), 674-693.

Mangolini, M., 1994. Apport de la fusion d'images satellitaires multicapteurs au niveau pixel en télédétection et photo-interprétation. Thèse de Doctorat en Sciences de l'Ingénieur, Université de Nice-Sophia Antipolis, France, 174 p.

Nishii, R., Kusanobu, S., Tanaka, S., 1996. Enhancement of low spatial resolution image based on high resolution bands. *IEEE Transactions on Geoscience and Remote Sensing* 34(5), 1151-1158.

Nuñez, J., Otazu, X., Fors, O., Prades, A., Palà, V., Arbiol, R., 1999. Multiresolution-based image fusion with additive wavelet decomposition. *IEEE Transactions on Geoscience and Remote Sensing* 37(3), 1204-1211.

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

Pradines, D., 1986. Improving SPOT image size and multispectral resolution. Proc. SPIE "Earth Remote Sensing using the Landsat Thematic Mapper and SPOT Systems", Vol. 660, pp. 78-102.

Price, J. C. 1999. Combining multispectral data of differing spatial resolution. *IEEE Transactions on Geoscience and Remote Sensing* 37(3), 1199-1203.

Ranchin, T., 1997. Wavelets, remote sensing and environmental modelling. In: Sydow, A. (Ed.), Proceedings, 15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics, Vol. 6, pp. 27-34.

Ranchin, T., Wald, L., 1996a. Benefits of fusion of high spatial and spectral resolutions images for urban mapping. In: Proceedings of the 26th International Symposium on Remote Sensing of Environment and the 18th Annual Symposium of the Canadian Remote Sensing Society, Vancouver, British Columbia, Canada, March 25-29, pp. 262-265.

Ranchin, T., Wald, L., 1996b. Merging SPOT-P and KVR-1000 for updating urban maps. In: Proceedings of the 26th International Symposium on Remote sensing of Environment and the 18th Annual Symposium of the Canadian Remote Sensing Society, Vancouver, British Columbia, Canada, March 25-29, pp. 401-404.

Ranchin, T., Wald, L., 1997. Fusion d'images HRV de SPOT panchromatique et multibande à l'aide de la méthode ARSIS: apports à la cartographie urbaine. In: Actes des journées scientifiques de Liège 1995, AUPELF-UREF, Montréal, Canada, pp. 283-290.

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

Ranchin, T., Wald, L., 2000b. Comparison of different algorithms for the improvement of the spatial resolution of images. In: Ranchin, T., Wald, L. (Eds.), Proceedings of the third conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images", Sophia Antipolis, France, January 26-28. SEE/URISCA, Nice, France, pp. 33-41.

Ranchin, T., Wald, L., Mangolini, M., 1994. 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. San Diego, California, USA, July 11-16, Vol. 2034, pp. 171-178.

Raptis, V. S., Vaughan, R. A., Ranchin, T., Wald, L., 1998. Assessment of different data fusion methods for the classification of an urban environment. In: Ranchin, T., Wald, L. (Eds.), Proceedings of the second conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images", Sophia Antipolis, France, January 28-30. SEE/URISCA, Nice, France, pp. 167-182.

Shensa, M.J., 1992. The discrete wavelet transform: wedding the a trous and Mallat algorithms. *IEEE Transactions on Signal Processing* 40(10), 2464-2482.

Terretaz, P., 1997. Comparison of different methods to merge SPOT P and XS data: Evaluation in an urban area. In: Gudmansen, P. (Ed.), Proceedings of 17th Symposium of EARSeL "Future trends in remote sensing", Lyngby, Denmark, 17-20 June. A. A. Balkema, Rotterdam, pp. 435-445.

Vaiopoulos, D., Nikolakopoulos, K., Skianis, G., 2001. A comparative study of resolution merge techniques and their efficiency in processing image of urban areas. In: Proceedings of the IEEE/ISPRS joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Rome, Italy, November, 8-9th, pp. 270-274.

Vrabel, J., 2000. Multispectral imagery advanced band sharpening study. *Photogrammetric Engineering & Remote Sensing* 66(1), 73-79.

Wald, L., 1998. A European proposal for terms of reference in data fusion. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXII, Part 7, pp. 651-654.

Wald, L., 1999. Some terms of reference in data fusion. *IEEE Transactions on Geoscience and Remote Sensing* 37(3), 1190-1193.

Wald, L., 2002. *Data fusion: definitions and architectures – fusion of images of different spatial resolutions*. Les Presses, Ecole des Mines de Paris, Paris, France, 200 p.

Wald, L., Baleynaud, J.-M., 1999. Observing air quality over the city of Nantes by means of Landsat thermal infrared data. *International Journal of Remote Sensing* 20(5), 947-959.

Wald, L., Ranchin, T., 2001. Data fusion for a better knowledge of urban areas. In: Proceedings of the IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Rome, Italy, November, 8-9th, pp. 127-132.

Wald, L., Ranchin, T, Mangolini, M., 1997. Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images. *Photogrammetric Engineering & Remote Sensing* 63(6), 691-699.

Woodcock, C. E., Strahler, A. H., 1987. The factor of scale in remote sensing. *Remote Sensing of Environment* 21(3), 311-332.

Yang, W., Cauneau, F., Paris, J.-P., Ranchin, T., 2000. Influence of landscape changes on the results of the fusion of P and XS images by different methods. In: Ranchin, T., Wald, L. (Eds.), Proceedings of the third conference "Fusion of Earth data: merging point

measurements, raster maps and remotely sensed images", Sophia Antipolis, France, January 26-28. SEE/URISCA, Nice, France, pp. 47-56.

Table captions

Table 1

Spectral bands, mean value and standard deviation of the panchromatic (PAN) and multispectral images (MS) (GV …grey values). Correlation coefficient between the original spectral bands and the PAN image resampled at 4 m

Table 2

Some statistics of the relative differences and the relative difference in variance (all in percent) and the correlation coefficient between the original and synthesised images for the spectral bands blue, green, red and NIR. See text for more explanations

Table 3

The error ERGAS for the various methods

Figure captions

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

Fig. 2. Pyramid representing the multiscale models.

Fig. 3. a) Region of the panchromatic image at 1 m ; b) region of the fused product in the NIR band synthesised by the method GLP-AABP at 1 m ; c) as b), but for the method UWT-RWM.

Fig. 4. Illustration of testing the second property in the NIR band at 4 m. Same geographical area as Fig. 3. From top to bottom: the original image, the GLP-AABP product and the UWT-RWM product.

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

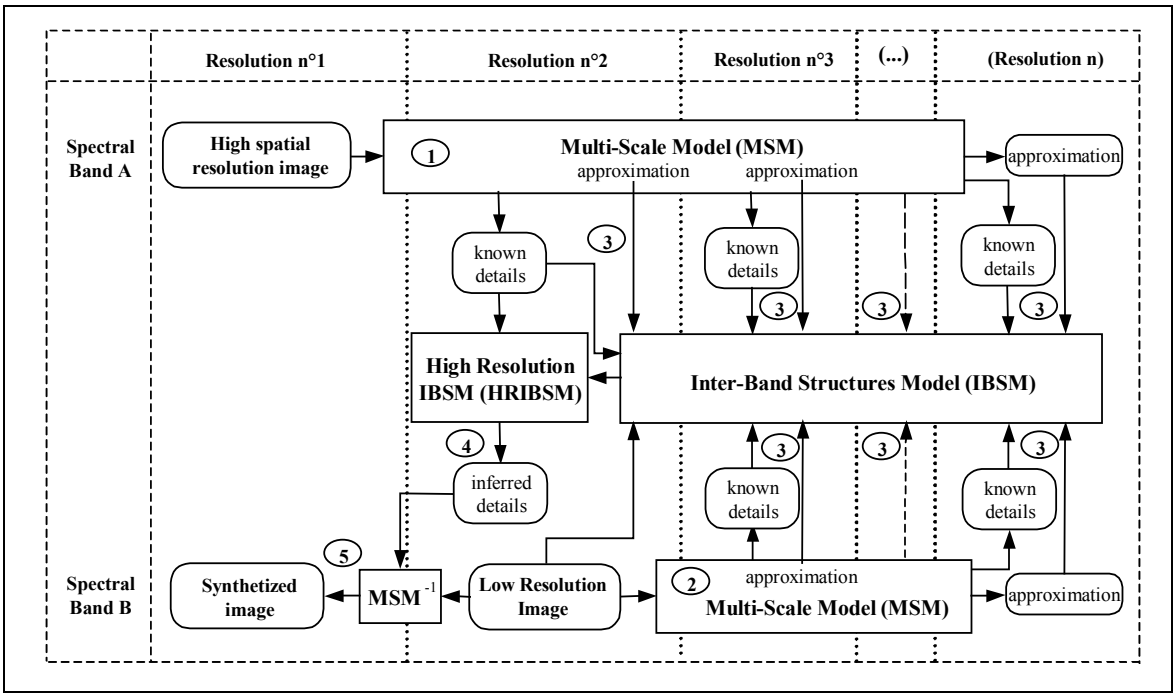

Fig. 1.

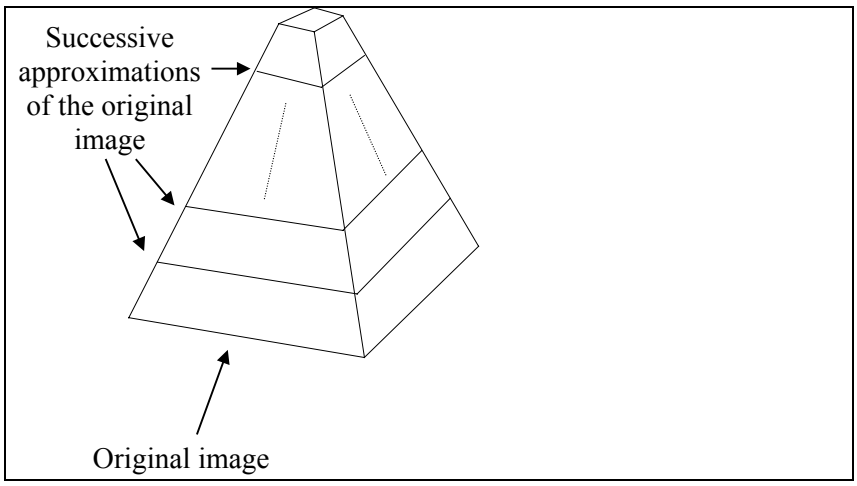
Successive
approximations
of the original
image
Original image

Fig. 2.

Fig. 3a.

Fig. 3b.

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

Fig. 3c.

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

Fig. 4.

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

	PAN	Blue	Green	Red	NIR
Spectral band (nm)	450 – 900	450 – 520	520 – 600	630 – 690	760 – 900
Mean value (GV)	352	328	339	249	467
Standard dev. (GV)	55	26	39	48	124
Corr. coeff.	1.00	0.271	0.409	0.323	0.875

Table 1

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

		Bias	Standard deviation	RMSE	Diff. in variance	Correlation coefficient
Blue	GLP-AABP	0.00	3	3	11	0.95
	UWT-AABP	-0.04	4	4	29	0.89
	UWT-RWM	0.01	3	3	23	0.91
Green	GLP-AABP	0.00	4	4	12	0.95
	UWT-AABP	-0.01	6	6	33	0.87
	UWT-RWM	0.01	5	5	27	0.91
Red	GLP-AABP	0.00	6	6	12	0.95
	UWT-AABP	-0.02	9	9	31	0.88
	UWT-RWM	0.04	8	8	25	0.91
NIR	GLP-AABP	0.00	7	7	10	0.97
	UWT-AABP	0.03	12	12	29	0.89
	UWT-RWM	0.03	8	8	20	0.96

Table 2

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

GLP-AABP	UWT-AABP	UWT-RWM
1.3	2.1	1.6

Table 3