Fusion of High Spatial and Spectral Resolution Images: The ARSIS Concept and Its Implementation

Thierry Ranchin and Lucien Wald

Abstract

In various applications of remote sensing, when high spatial resolution is required in addition with classification results, sensor fusion is a solution. From a set of images with different spatial and spectral resolutions, the aim is to synthesize images with the highest spatial resolution available in the set and with an appropriate spectral content. Several sensor fusion methods exist; most of them improve the spatial resolution but provide poor quality of the spectral content of the resulting image. Based on a multiresolution modeling of the information, the ARSIS concept (from its French name "Amélioration de la Résolution Spatiale par Injection de Structures") was designed with the aim of improving the spatial resolution together with a high quality in the spectral content of the synthesized images. The general case for the application of this concept is described. A quantitative comparison of all presented methods is achieved for a SPOT image. Another example of the fusion of SPOT XS (20-m) and KVR-1000 (2-m) images is given. Practical information for the implementation of the wavelet transform, the multiresolution analysis, and the ARSIS concept by practitioners is given with particular relevance to SPOT and Landsat imagery.

Introduction

In various applications, the benefit of obtaining multispectral images with the highest spatial resolution available has been demonstrated, particularly for vegetation, land-use, precision farming, and urban studies. On the one hand, high spatial resolution is necessary for an accurate description of shapes, features, and structures. On the other hand, depending on the application and the level of land-cover complexity, different types of land use are better classified if high spectral resolution images are used. Hence, there is a desire to combine both high spatial and high spectral resolution images with the aim of obtaining the most complete and accurate (in terms of spectral band) description of the observed area.

Several approaches for sensor fusion exist which have been applied to data sets consisting of multispectral images at a low spatial resolution and images at a higher spatial resolution but with a lower spectral content. Examples of such a data set are the SPOT-XS (three bands, 20-m) and SPOT-P (panchromatic, 10-m) images. These methods aim at constructing synthetic multispectral images having the highest spatial resolution available within the data set (e.g., the three XS bands at 10 m in the case of SPOT). This paper discusses some of these methods and their advantages and disadvantages, and underlines the potential of these methods based on multiresolution analysis and wavelet transform. We are concerned only with those methods which claim to provide synthetic images close to reality when enhancing spatial resolution, and not those which provide only a better visual representation of the data set (e.g., Carper *et al.*, 1990; Mangolini, 1994). Many articles have been published which demonstrate that the spectral content of an image changes as the spatial resolution changes (see, e.g., an extensive discussion in Wald *et al.* (1997)). Note that several authors dealing with the methods under consideration (including the present authors) improperly write "preservation of spectral content," an inappropriate shortcoming for "high-quality transformation of the multispectral content when increasing the spatial resolution."

The currently most used methods are presented in the first section. The need for advanced methods is demonstrated for high-quality synthesis, using mathematical tools such as the wavelet transform, presented in the second section, and multiresolution analysis, discussed in the third section. Skilled practitioners can implement these tools using the practical information given in the fourth section. The ARSIS concept is described in the following section. The case of application of this concept to SPOT and Landsat imagery is detailed. A quantitative comparison of all presented methods is carried out for a SPOT image using the approach proposed by Wald *et al.* (1997). The potential of the ARSIS concept is further demonstrated through an example of fusing SPOT XS and KVR-1000 (panchromatic, 2-m) images.

Background - The Brovey, IHS, and PCA Methods

As a first requirement for these methods, all images should be superimposed onto each other, once all images have been set to the lowest available spatial resolution (e.g., 20 m in the case of SPOT). If not already done by the observation system (e.g., Landsat) or by the data provider, this can be done by means of standard methods available in public or commercial software packages for image processing. Blanc *et al.* (1998a) and Wald *et al.* (1997) discussed the influences of, respectively, the quality of the co-registration and the resampling operator on the final results. The relative discrepancies between the results are a few percent; these influences can be kept very small provided the co-registration is accurate enough and the operator is appropriate enough. In the following, for the sake of simplicity, the term "image of lowest resolution" will denote the projected resampled image of lowest resolution, if this is required.

Ecole des Mines de Paris, Groupe Télédétection & Modélisation, BP 207, 06904 Sophia Antipolis cedex, France (thierry.ranchin@cenerg.cma.fr).

Photogrammetric Engineering & Remote Sensing
Vol. 66, No. 1, January 2000, pp. 49–61.

0099-1112/00/6601–000$3.00/0

Three methods are commonly used: the Brovey transform, the IHS (Intensity, Hue, Saturation) method, and the PCA (Principal Component Analysis) method. The Brovey transform is based on spectral modeling, while the IHS and PCA methods call upon projection techniques. They apply to the spectral images B_i resampled at the highest spatial resolution. These methods are available in most commercial software packages for satellite image processing, which may explain their large use by practitioners. Here the ERDAS Imagine software was used for their application. Details on these methods can be found in, e.g., Carper *et al.* (1990), Chavez *et al.* (1991), or Pohl and Van Genderen (1998).

Let A_h be the high spatial resolution image, B_l the multispectral image, h the original spatial resolution of A, and l the original spatial resolution of B ($l < h$). The Brovey transform applies to the digital counts of three spectral bands B'_{ih} ($i = 1, 2, 3$), where B'_{ih} is the image B_{il} resampled at resolution h, and to an image A_h of a better resolution, although it should deal with radiances and the software should request the calibration coefficients. An example is the SPOT case, where B'_{ih} are the XS1, XS2, and XS3 bands at original resolution of 20 m, resampled to 10 m, and A_h is the P band with a spatial resolution of 10 m. The synthetic bands B^*_{ih} are given by

$$B^*_{ih} = B'_{ih}\, A_h / (B'_{1h} + B'_{2h} + B'_{3h}). \tag{1}$$

For the SPOT case, it is far better to use the CNES method, called P + XS (Anonymous, 1986), as shown later. The Brovey transform assumes that the spectral range of the A_h image covers the spectral range of the sum of the B_{il}. This is not true, e.g., for the SPOT-XS3 band, which lies outside the P band. This assumption causes a spectral distortion. It should be noted that this assumption is also made by the IHS and PCA methods. Finally, not using radiances in the Brovey transform induces a bias in synthesized images.

The method "relative spectral contribution," which applies to radiances, generalizes the Brovey transform and the P + XS method. It restrains itself to the spectral bands B_i lying within the spectral range of the A_h image (see, e.g., Wiemker *et al.* (1998)). The sum $(B'_{1h} + B'_{2h} + B'_{3h})$ in Equation 1 is replaced by the sum of all these spectral bands B_i. The method does not tell what to do when B_i lies outside the spectral range of A_h; the P + XS recommends a simple duplication for the XS3 band, which lies outside the P band. For this method, the P + XS method, or the Brovey transform, a Fourier transform of Equation 1 shows that there is an influence of the other spectral bands on the assessment of B^*_{ih}, thus inducing a spectral distortion (which adds to those already discussed for the Brovey transform). This influence may range from low to high, depending mostly upon the landscape (see a discussion in Wald *et al.* (1997)) and also on the modulation transfer functions of the sensors. If the representation for an object is fairly close in the different spectral bands, the influence will be low. This method cannot resolve local anti-correlations between spectral bands with a high accuracy in the synthesizing of the spectral content. It does not even preserve the original spectral content once the synthesized images B^*_{ih} are brought back to the original low spatial resolution.

In the IHS method, each of the three bands B_i are labeled as blue, green, and red, respectively. Then, these color components are converted into intensity (I), hue (H), and saturation (S) components using, for example, the model for colors of "Commission Internationale pour l'Eclairage." The next step is the substitution of the intensity by the high spatial resolution image A_h. Refinements can be made which include the substitution of a linear combination of the A_h values and the original intensity. The last step performs the inverse model, converting IHS components into blue, green, and red components, which are the searched synthetic images B^*_{ih}. The method can apply to either digital counts or to radiances. In any case, the dynamics of the signal in each bands, including A_h should be adjusted in order to make them similar. This may cause a distortion of the spectral content.

The PCA method is rather similar in essence to the IHS method but applies to two or more spectral bands. The first component (which acts as the intensity in the IHS method) can be replaced by the high spatial resolution image A_h (or by a combination of this image and of the first component). An inverse PCA transformation allows one to synthesize the multispectral high spatial resolution images B^*_{ih}.

The merging methods under consideration aim at constructing synthetic images B^*_h close to reality. Wald *et al.* (1997) established the properties of such synthetic images as follows:

- 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 spatial resolution h; and
- 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.

The Brovey, IHS, and PCA methods are not suitable in this aim. They clearly do not respect any of these properties, even the first one, because of their very construction. A new sensor fusion concept is proposed, allowing the improvement of the spatial resolution of images up to the best available in the set of images and the quality of the transformation of the spectral content when increasing the spatial resolution. This concept is called ARSIS, after the acronym of its French name "amélioration de la résolution spatiale par injection de structures" (improvement of spatial resolution by structure injection). It has been designed in a generic way as a concept should be, in order to meet as close as possible the above-mentioned properties, and transcending the mathematical tools used for its implementation. The approach itself is not new as discussed below, but it has never been expressed as a concept before the joint work of Ecole des Mines de Paris and Aérospatiale (Mangolini *et al.*, 1992; Ranchin, 1993). Though it has been intensively used by its authors, and though several other scientists have published studies using the concept without naming it, it is the first time that this concept is fully explained and that detailed information is given for its practical implementation and further use by practitioners, software companies, or data/products providers.

The ARSIS concept makes use of a multiscale method for the description and the modeling of the missing information between the images A_h and B_l. The multiscale method, used mostly for its implementation, is the multiresolution analysis, together with the wavelet transform. Other tools exist: Blanc *et al.* (1998b) used iterated filter banks instead of the wavelet transform, Tom (1987) used Gaussian filters, and Chavez *et al.* (1991) used self-defined filters in their HPF method (see later). The next sections present the wavelet transform, the multiresolution analysis, and the practical implementation of both tools.

The Wavelet Transform

As with the Fourier transform, the wavelet transform performs a decomposition of the signal on a base of elementary functions: the wavelets. The base is generated by dilations and translations of a single function ψ called the mother wavelet: i.e.,

$$\psi_{a,b} = |a|^{-1/2}\, \psi[(x - b)/a] \tag{2}$$

where $a, b \in \Re$, and $a \neq 0$. a is called the dilation step and b the translation step. Many mother wavelets exist. They are all oscillating functions, and are well localized both in time and

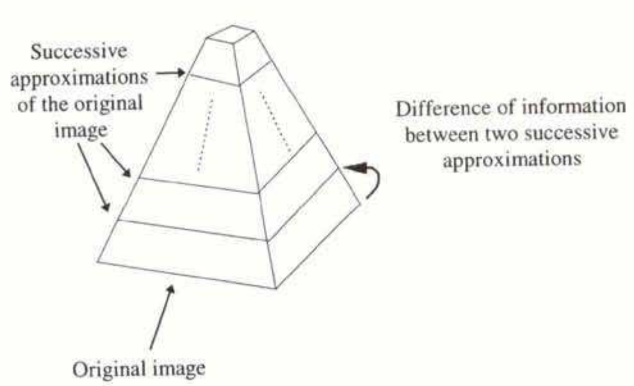

Figure 1. Pyramid representing the multiresolution analysis combined with the wavelet transform.

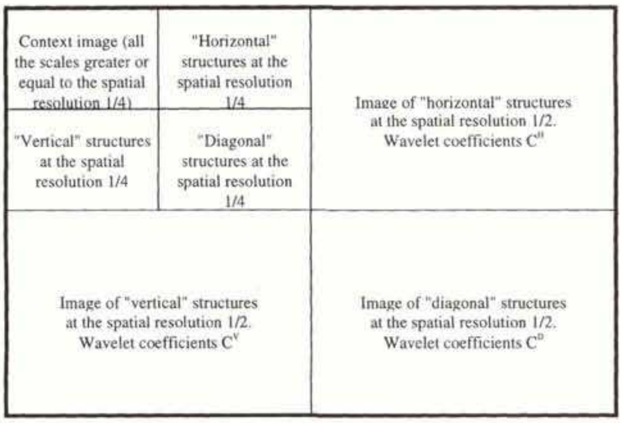

Figure 2. Presentation of a multiresolution analysis using the Mallat algorithm. Original resolution of the image is 1.

frequency. All the wavelets have common properties such as regularity, oscillation, and localization, and they satisfy an admissibility condition. For more details about the properties of wavelets, one can refer to Meyer (1990) or Daubechies (1992). Even if they have commons properties, each of them leads to a unique decomposition of the signal related to the selected mother wavelet. In the one-dimensional case, the continuous wavelet transform of a function $f(x)$ is

$$WT_f(a, b) = \langle f, \psi_{a,b}\rangle = \frac{1}{\sqrt{|a|}}\int_{-\infty}^{+\infty} f(x)\,\overline{\psi\left(\frac{x-b}{a}\right)}\,dx \qquad (3)$$

where $\overline{\psi\left(\frac{x-b}{a}\right)}$ is the complex conjugate of ψ. The computation of the wavelet transform for each scale a and each location b of a signal $f(x)$ provides a local representation of $f(x)$, and the information content is represented by the wavelet coefficient $WT_f(a,b)$. The process can be reversed and the original signal reconstructed exactly (without any loss) from the wavelet coefficients by

$$f(x) = \frac{1}{C_\psi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} WT_f(a, b)\psi_{a,b}(x)\,\frac{da\,db}{a^2} \qquad (4)$$

where C_ψ is the admissibility condition of the mother wavelet.

Discrete versions of the wavelet transform exist and are applied to signals filters.

Multiresolution Analysis

Figure 1 is a very convenient description of multiresolution analysis and more generally of pyramidal algorithms (Mallat, 1989). The base 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 successive approximations have coarser and coarser spatial resolutions. The computation of the approximations is done using a base of functions, called the scale functions. The base is generated following the same scheme as the one used for the generation of the wavelet base. The base of the pyramid is the landscape measured by the sensor. In this scheme, the use of the 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) by wavelet coefficients. If the process of multiresolution analysis is inverted, the original image can be exactly reconstructed from one approximation and from the different wavelet coefficients describing the differences in signal between this approximation and the original image: this is called synthesis.

As we are processing images, the wavelet and the scale functions are applied first in columns and then in lines (rows). This leads to a representation of the information using the scheme proposed in Figure 2 (taken from Ranchin and Wald (1993)). The dilation of both the wavelet and the scale function is obtained by the sub-sampling of the original image. Hence, if the original image comprises, e.g., 768 lines by 1024 columns, the first approximation is 384 lines by 512 columns, as well as the three wavelet coefficients images. The successive approximations of the images are also called context images. The first context image contains all the scales greater than half the original spatial resolution (1/2 in Figure 2, i.e., greater than 20 m for SPOT P). The three wavelets coefficients images represent the structures with sizes comprised between the original spatial resolution and half this resolution (i.e., between 10 and 20 m for P) for the diagonal (C^D), vertical (C^V), and horizontal (C^H) directions. In the second context image, all the scales greater than a quarter of the original resolution are represented (i.e., 40 m for P), and the wavelet coefficient images contain the scales between half and a quarter of the original resolution (i.e., 20 and 40 m for P).

Practical Implementation of the Multiresolution Analysis and Wavelet Transform

Mallat's algorithm can be implemented using a filter bank structure (Figure 3). We detail here its optimized implementation for the ARSIS concept. In this scheme, $f_j(x, y)$ represents the original image, where x is the column (beginning from 1), and y is the line or row (beginning from 1). In the case of the practical implementation for SPOT imagery, H and G are the four-tap filters designed by Daubechies (1988) and discussed later. The columns and rows are processed separately. Filter H is applied to the columns of $f_j(x, y)$, and the same is done for filter G. Both resulting images are re-sampled (operation $\downarrow 2$ in Figure 3): one column over two is removed. Then on each re-sampled image filters H and G are applied. The resulting four images are re-sampled: one row over two is removed. This results in four images:

- $f_{j+1}(x, y)$ is the approximation (context) with half the spatial resolution of the original $f_j(x, y)$ image one; and
- the three wavelet coefficients images $C^D_{j+1}(x, y)$, $C^V_{j+1}(x, y)$, and $C^H_{j+1}(x, y)$.

Table 1 gives the coefficients of the four-tap filter H defined by Daubechies. All these coefficients have to be divided by $\sqrt{2}$ for normalization purposes. The filters are applied as shown in

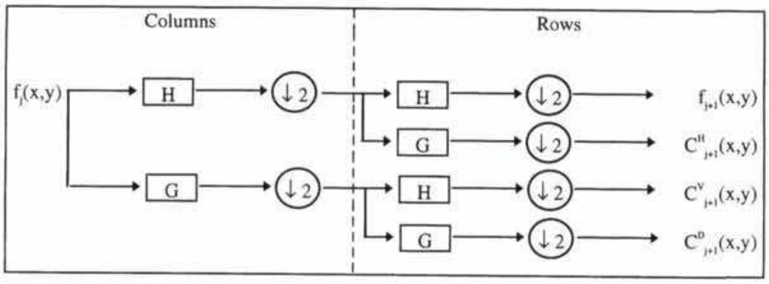

Figure 3. Implementation of the analysis of the Mallat's algorithm into a filter bank structure.

TABLE 1. VALUES OF THE COEFFICIENTS OF THE FILTER FOR THE WAVELET DEFINED BY DAUBECHIES (1988)

H(0)	H(1)	H(2)	H(3)
0.482962913145	0.836516303738	0.224143868042	−0.129409522551

Figure 4, which presents their application along a row. The new value for the current pixel (x, y) is computed as a multiplication between the coefficients of the filters and the pixels: i.e.,

$$\text{new value} = H(3)\, f(x-2, y) + H(2)\, f(x-1, y) + H(1)\, f(x, y) + H(0)\, f(x+1, y) \qquad (5)$$

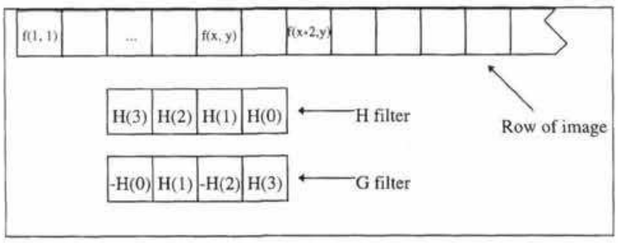

Figure 4. Position of the filters in the analysis for the application of ARSIS to SPOT imagery. In this figure, $f(x,y)$ denotes the function on which the filter is applied, e.g. $f_j\,(x, y)$. The column x is odd. Standard solutions can be adopted for the borders.

A similar equation is applied, but for filter G, which is derived from the H filter (see Figure 4). Then, due to the subsampling, the next pixel to be computed is $(x + 2, y)$. The new value of pixel $(x + 1, y)$ is temporarily set to 0. Actually, this pixel is not processed at all because it will be removed by subsampling (see Figure 3). Once all of the image is processed for the columns, the same process is applied to all rows along the column direction.

From the approximation $f_{j+1}\,(x, y)$ and from the three wavelet coefficients images $C^D_{j+1}\,(x, y)$, $C^V_{j+1}\,(x, y)$, and $C^H_{j+1}\,(x, y)$, one can exactly reconstruct the original image $f_j(x, y)$. In the synthesis, an over-sampling $\uparrow 2$ is necessary (Figure 5). It is obtained by adding a zero between the pixels (Figure 6). In the case of orthogonal filters, $\tilde{H}$ and $\tilde{G}$ are the same filters as those used in the analysis (i. e., H and G). First, an over-sampling in columns is applied to the approximation and to the three coefficients images. Then, either the filter H or G is applied to each pixel, including those set to zero, and according to the scheme in Figure 6. Results are summed two by two as shown in Figure 5. An over-sampling is applied in lines, prior to the application of filter H and G. The final summation provides the original $f_j(x, y)$ (or synthesized) image after a multiplication by 4.

It is recommended that the implementation of Mallat's algorithm be checked. The following scheme can be employed, given any image. Apply analysis (Figure 3) with one iteration: the first approximation $f_{j+1}\,(x, y)$ is obtained. Then perform

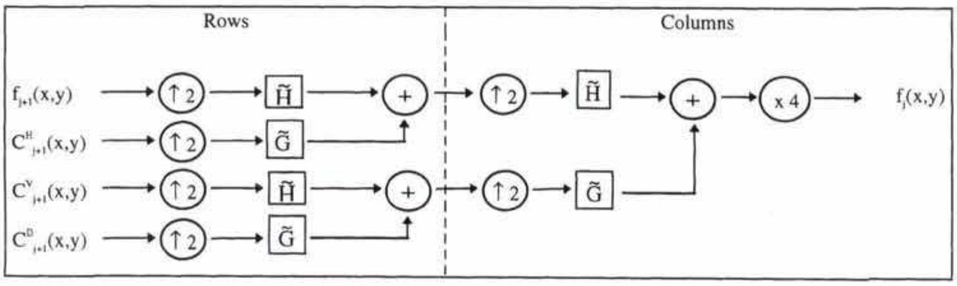

Figure 5. Implementation of the synthesis of the Mallat's algorithm into a filter bank structure.

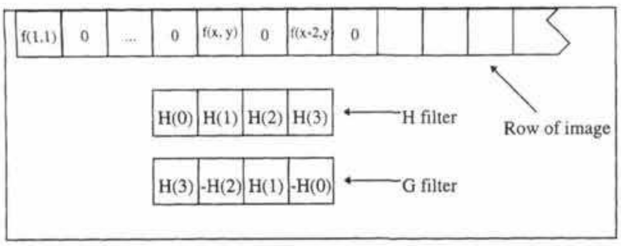

Figure 6. Position of the filters for the synthesis in the application of ARSIS to SPOT imagery. In this figure, $f(x, y)$ denotes the function on which the filter is applied, i.e., $f_{j+1}(x, y)$, $C^H_{j+1}(x, y)$, $C^V_{j+1}(x, y)$, or $C^D_{j+1}(x, y)$. The column x is odd. Standard solutions can be adopted for borders of image.

synthesis (Figure 5) on $f_{j+1}(x, y)$. The resulting image should be identical to the original, except for the borders of the image. This can be checked by computing the difference between both images, pixel by pixel. The whole checking procedure should be performed for more than one iteration.

The ARSIS Concept

The multiscale method is applied to the two images A and B. A scale-by-scale description of the information content of both images is obtained. The missing spatial information (high frequency) between A_h and B_l, represented by the wavelet coefficients, is extracted and used for the construction of the synthetic image B^*_h. Figure 7 illustrates this approach in the case of a pyramidal description. The missing information to be injected in pyramid B from image A is located in the missing bottom of pyramid B (dotted line). Only this part is needed to improve the spatial resolution of image B. But if the missing information is set equal to that provided by image A, the synthesized image B^*_h will not be equivalent to "what would be seen by sensor B if it has the spatial resolution of sensor A." Hence, in order to improve the quality of the synthesized image, a transformation should be applied to convert the information provided by the multiscale representation of image A into the information needed for the synthesis of image B.

Examples of the application of this concept have been given by Tom (1987), Mangolini *et al.* (1992; 1993), Ranchin *et al.* (1994; 1996), Iverson and Lersch (1994), Garguet-Duport (1994; 1996), Li *et al.* (1995), Yocky (1996), Ranchin and Wald (1998), Blanc *et al.* (1998b), and Zhou *et al.* (1998). They are all based on a multiresolution pyramidal approach. The HPF method (Chavez *et al.* 1991) and the methods of Pradines (1986) and Guo and Moore (1998) are other examples of the ARSIS concept, though not based on a pyramidal description. In the HPF method, a high pass filtering (HPF) is applied to the high spatial resolution image A_h in order to extract the high frequencies representing the small structures between scales h and l. Then, these high frequencies are introduced in the multispectral image B_l by addition, which leads to the synthetic image B^*_h. In the Pradines method, the relative spatial distribution of the high-resolution signal is injected into the low-resolution spectral image for each low-resolution pixel. In the SPOT case, the 20-m XS pixel is shared in four 10-m pixels using the relative distribution observed in the P image for these four pixels. Guo and Moore (1998) discuss a similar method, called "pixel block intensity modulation" (PBIM), and give an example using Landsat TM imagery. In these HPF, Pradines, and PBIM methods, the differences in the representation of the high frequency information in the spectral bands are not taken into account.

Figure 8 presents the general scheme for the application of the ARSIS concept. First, a multiresolution analysis using the wavelet transform is used to compute the wavelet coefficients and the approximations of image A (Step 1 in Figure 8). The same operation is applied to image B (Step 2). The wavelet coefficients provided by each decomposition are used to adjust the parameters of a model of transformation of the known wavelet coefficients of image A into the known wavelet coefficients of image B (Step 3). From this model is derived another one which converts the known wavelet coefficients of image A into the inferred (computed, Step 4) wavelet coefficient of image B. Finally, the inversion of the multiresolution analysis (WT^{-1}) allows the synthesis of the image B^*_h with the spatial resolution of image A (Step 5).

Figure 9 details the application of the ARSIS concept to the case of SPOT imagery, which may also be applied to Landsat and future missions. The set of images is composed of a panchromatic image at the spatial resolution of 10 m and three multispectral images XS1, XS2, and XS3 at the spatial resolution of 20 m. The process is applied to each XS*i* image separately. Two iterations of the multiresolution analysis using the wavelet transform are applied to the original panchromatic (P) image and one iteration to the original XS*i* image. Models are computed for the transformation of each panchromatic wavelet coefficient image C^D_{P20-40}, C^V_{P20-40}, and C^H_{P20-40} into each XS*i* wavelet coefficient image $C^D_{XS20-40}$, $C^V_{XS20-40}$, and $C^H_{XS20-40}$ (see proposed models below). Then, these models are applied to the wavelet coefficient images C^D_{P10-20}, C^V_{P10-20}, and C^H_{P10-20}, for the computation of the missing wavelet coefficient images $C^D_{XS10-20}$, $C^V_{XS10-20}$, and $C^H_{XS10-20}$. Finally, the synthesis step reconstructs the high spatial resolution XSi image (XSi-HR).

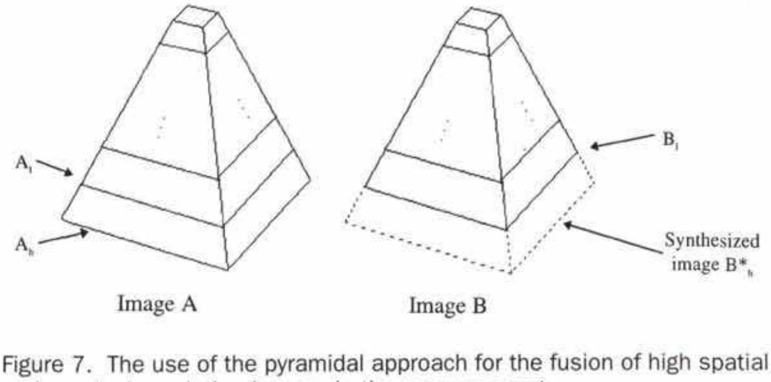

Figure 7. The use of the pyramidal approach for the fusion of high spatial and spectral resolution images in the ARSIS concept.

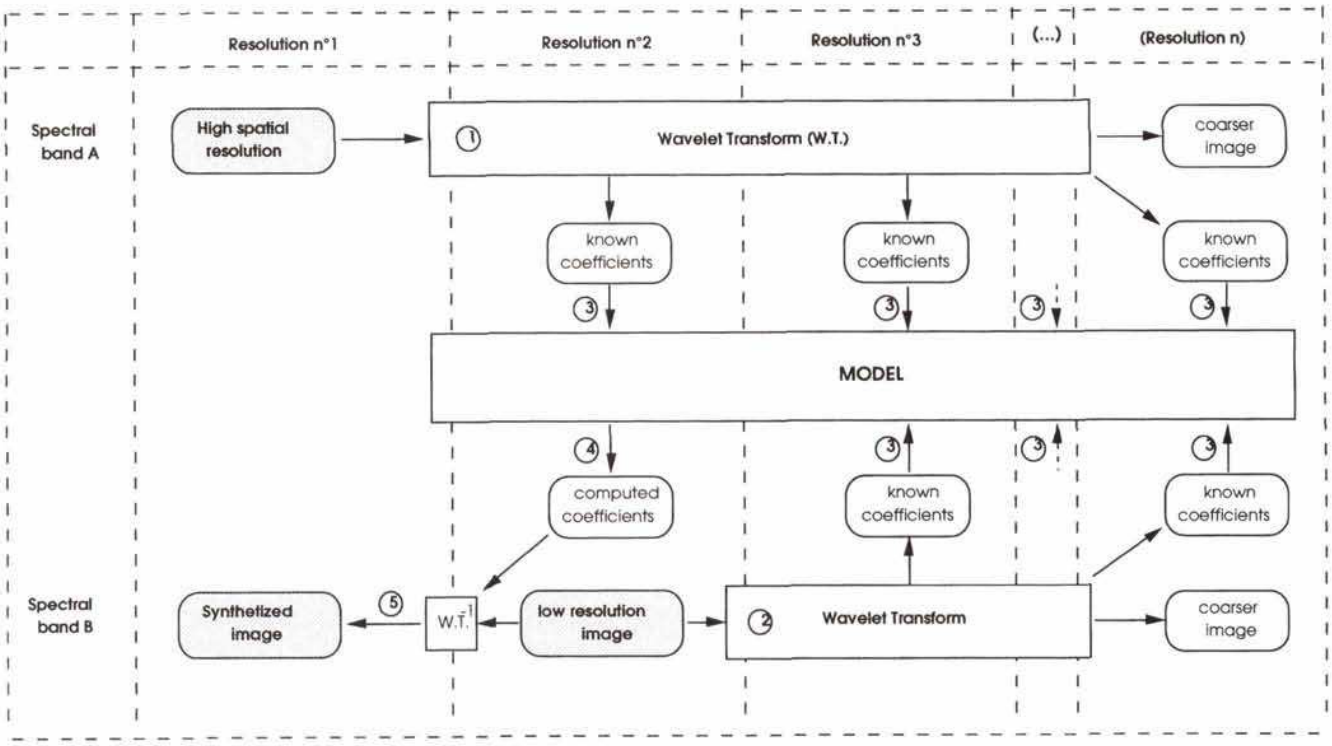

Figure 8. General scheme for the application of the ARSIS concept using wavelet transform (WT) and inverse wavelet transform (WT^{-1}). See text for further comments

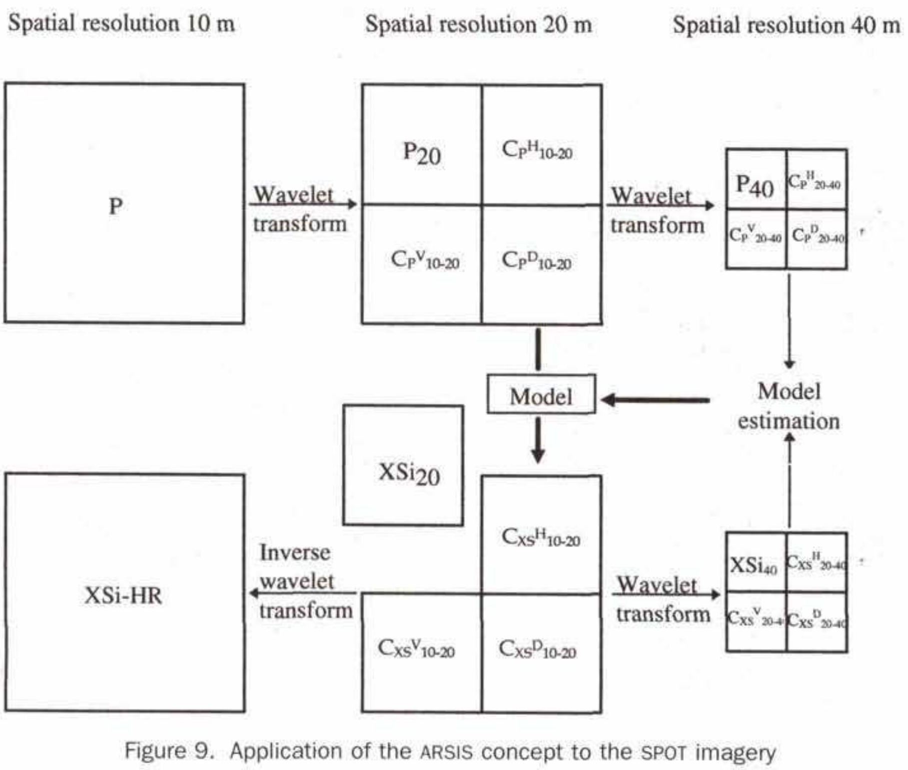

Figure 9. Application of the ARSIS concept to the SPOT imagery

The model is obviously a key point for an efficient synthesis. Many models can be proposed. For example, the model used by Ranchin *et al.* (1994; 1996) takes into account the physics of both images and the correlation or anti-correlation existing between the wavelet coefficients images. Iverson and Lersch (1994) used a model based upon neural networks. The model can have various forms and may take into account more than one scale. The simplest model (Model 1) is the identity model proposed by Mangolini *et al.* (1992) and Yocky (1996): i.e.,

$$C^{Z}_{\mathrm{XS}10-20} = C^{Z}_{\mathrm{P}10-20} \text{ for } Z = D,\ V, \text{ or } H. \tag{8}$$

This model is that which is implicitly used in the HPF method and in Tom's method (Tom, 1987). Model 1 does not take into account the spectral differences in high-frequency information between panchromatic image P and spectral bands XS and gives poor results. This can be seen directly from the equations but also from experiments (Mangolini *et al.*, 1992). The following models are more accurate.

Model 2 is based on the adjustment of the means and variances of the wavelet coefficients images computed between 20 and 40 m as proposed by Mangolini *et al.* (1992): i.e.,

$$C^{Z}_{\mathrm{XS}10-20} = a^{Z}\, C^{Z}_{\mathrm{P}10-20} + b^{Z} \text{ for } Z = D,\ V, \text{ or } H. \tag{9}$$

This model denotes the means of $C^{Z}_{\mathrm{XS}20-40}$ by m^{Z} (XS), respectively, and the standard deviation of $C^{Z}_{\mathrm{XS}20-40}$, by σ^{Z} (XS). Similar notations hold for the P coefficients. The following quantities are computed:

$$a^{Z} = \sigma^{Z}(\mathrm{XS})/\sigma^{Z}(\mathrm{P}) \text{ and}$$
$$b^{Z} = m^{Z}(\mathrm{XS}) - a^{Z}\, m^{Z}(\mathrm{P}) \text{ for}$$
$$Z = D,\ V, \text{ or } H.$$

In the case of the fusion of images from Landsat TM band 6 (120 m) and, e.g., Landsat TM band 4 (30 m), as was done by Wald and Baleynaud (1999), two wavelet coefficients images need to be synthesized: i.e., between 120 and 60 m, and between 60 and 30 m. Then, the equations for Model 2 are

$$C^{Z}_{(\mathrm{TM6})60} = a^{Z}\, C^{Z}_{(\mathrm{TM4})60} + b^{Z} \tag{10}$$
$$C^{Z}_{(\mathrm{TM6})30} = a^{Z}\, C^{Z}_{(\mathrm{TM4})30} + b^{Z} \text{ for } Z = D,\ V, \text{ or } H$$

where $C^{Z}_{(\mathrm{TM6})120}$, $C^{Z}_{(\mathrm{TM6})60}$, $C^{Z}_{(\mathrm{TM6})30}$, $C^{Z}_{(\mathrm{TM4})120}$, $C^{Z}_{(\mathrm{TM4})60}$ and $C^{Z}_{(\mathrm{TM4})30}$ denote the sets of wavelet coefficients (D, V, H) as above but for the Landsat-TM; m^{Z} (TM6) and m^{Z} (TM4) are the means of $C^{Z}_{(\mathrm{TM6})120}$ and $C^{Z}_{(\mathrm{TM4})120}$, respectively; σ^{Z} (TM6) and σ^{Z} (TM4) are the standard deviations of $C^{Z}_{(\mathrm{TM6})120}$ and $C^{Z}_{(\mathrm{TM4})120}$, respectively;

$$a^{Z} = s^{Z}(\mathrm{TM6})/s^{Z}(\mathrm{TM4}); \text{ and}$$
$$b^{Z} = m^{Z}(\mathrm{TM6}) - a^{Z}\, m^{Z}(\mathrm{TM4}).$$

A hybrid Model 2 was used by Garguet-Duport *et al.* (1994; 1996), Li *et al.* (1995), and Zhou *et al.* (1998) where SPOT-P was stretched to have the same mean and variance as the TM bands. Because the mean of any wavelet coefficient image (m^{Z}) is null (less than 10^{-3}), and the wavelet transform is linear, the results of the Model 2 and hybrid Model 2 are very similar.

Model 3 is also based on Equation 9 but a^{Z} and b^{Z} are now computed using an adjustment between $C^{Z}_{\mathrm{XS}20-40}$ and $C^{Z}_{\mathrm{P}20-40}$ using either least-squares fitting or axis of inertia (Mangolini *et al.*, 1992). Several experiments showed that both adjustment methods provide similar results. For SPOT, several tens of application cases made at the Ecole des Mines de Paris showed that Model 3 gives better results than Model 2 (unpublished results).

Because the physics are taken into account, the ARSIS concept should apply to radiances. However, if one uses a linear model such as Models 1, 2, or 3, and if the calibration law is linear as for SPOT or Landsat, identical results are obtained using digital counts directly.

The choice of the wavelet transform used in the multiresolution analysis is also of importance. It has been discussed by Ranchin and Wald (1997a). The shortest orthogonal filters lead to the best results. In the case of SPOT and Landsat imagery, Mallat's algorithm is used with the four-tap filter proposed by Daubechies, as shown in the previous section.

Comparison of the Methods

How to assess the quality of a merging method has been discussed by Wald *et al.* (1997), a work distinguished by the ASPRS. A formal approach was established and was followed to assess the quality of the presented methods: Brovey, IHS, PCA, and ARSIS Models 1 and 2 (these latter two models were the easiest to implement). The methods were applied to the same SPOT image of Barcelona, Spain as in the article by Wald *et al.* (1997), who discussed this image and the geographical area. This allowed a further comparison with the "duplication," "P + XS" from CNES, and ARSIS-RWM methods (another model established by Ranchin *et al.*, (1994)). The rationale to use images of urban areas was given by Wald *et al.* (1997). The conclusions drawn from this example have been validated for several other cases by the authors and also by other authors (Terretaz, 1997; Raptis *et al.*, 1998; Wiemker *et al.*, 1998; Zhou *et al.*, 1998).

The first step is the visual inspection of the synthesized images, which will show the major drawbacks of a method. Once the contrast table is adjusted for each B^{*}_{h}, one can compare them. The different B^{*}_{h} images are very close to one another for all methods, except for the IHS method, which produces an image of poor quality in the XS3 band.

The first property "any synthetic image B^{*}_{h}, once degraded to its original resolution l, should have is that it be as identical as possible to the original image B_{l}"; is this clearly not satisfied by the Brovey, IHS, and PCA methods as discussed earlier in an analytical way (nor by the P + XS). The influence of A_{h} and the other spectral bands B_{l} in the synthesized image B^{*}_{h} does not disappear when reducing the resolution to 20 m. The methods built within the ARSIS concept as well as the duplication or interpolation methods are inherently built to satisfy this first property, with reservations regarding the degradation process as discussed by Wald *et al.* (1997).

To test the second and third properties, the P and XS images were degraded to a resolution of 20 and 40 m, respectively. Then, the images were synthesized at a 20-m resolution and compared to the original XS images. Tables 2, 3, and 4 give, respectively, the first set of criteria for the XS1, XS2, and XS3 images that provide a global view of the discrepancies between the original image and the synthetic image. It contains

- The bias, and its value relative to the means of the original image. It is the difference between the means of the original image and of the synthetic image.
- The difference in variances, as well as its value relative to the variance of the original image. It expresses the quantity of information lost or added during the enhancement of the spatial resolution.
- The correlation coefficient between the original and synthetic image. It shows the similarity in small size structures between the original and synthetic images.
- The standard deviation of the difference image, as well as its value relative to the mean of the original image. It globally indicates the level of error at any pixel.

The Brovey transform gives very poor results in all aspects of Tables 2 through 4, due to the bias and the spectral distortion it induces. The IHS method performs better, but the results remain of poor quality: the bias is high, not enough details are injected

TABLE 2. SOME STATISTICS ON THE DIFFERENCES BETWEEN THE ORIGINAL AND SYNTHESIZED IMAGES, IN RADIANCE (Wm^{-2} st^{-1} μm^{-1}) OR RELATIVE VALUE, FOR THE XS1 BAND. VALUES FOR ARSIS-RWM ARE TAKEN FROM WALD *ET AL.* (1997).

	Brovey	IHS	PCA	ARSIS Model 1	ARSIS Model 2	ARSIS RWM
Bias (ideal value: 0)	36.90	−5.85	−2.13	0.00	0.00	0.00
relative to the mean XS value	64%	−10%	−4%	0%	0%	0%
Actual variance-estimate (ideal value: 0)	100	32	−67	−6	−4	7
relative to the actual variance	70%	22%	−47%	−4%	−3%	5%
Correlation coefficient between XS and estimate (ideal value: 1)	0.97	0.92	0.98	0.98	0.99	0.99
Standard-deviation of the differences (ideal value: 0)	5.9	4.8	3.8	2.2	2.1	1.9
relative to the mean of XS value	10%	8%	7%	4%	4%	3%

TABLE 3. AS IN TABLE 2, BUT FOR THE XS2 BAND

	Brovey	IHS	PCA	ARSIS Model 1	ARSIS Model 2	ARSIS RWM
Bias (ideal value: 0)	30.60	−4.58	−2.81	0.00	0.00	0.00
relative to the mean XS value	64%	−10%	−6%	0%	0%	0%
Actual variance-estimate (ideal value: 0)	172	31	−113	−11	−8	7
relative to the actual variance	77%	14%	−51%	−5%	−3%	3%
Correlation coefficient between XS and estimate (ideal value: 1)	0.98	0.96	0.98	0.99	0.99	0.99
Standard-deviation of the differences (ideal value: 0)	8.1	4.0	4.9	2.6	2.3	1.9
relative to the mean of XS value	17%	8%	10%	5%	5%	4%

TABLE 4. AS IN TABLE 2, BUT FOR THE XS3 BAND

	Brovey	IHS	PCA	ARSIS Model 1	ARSIS Model 2	ARSIS RWM
Bias (ideal value: 0)	35.50	−5.49	−0.42	0.00	0.00	0.00
relative to the mean XS value	65%	−10%	1%	0%	0%	0%
Actual variance-estimate (ideal value: 0)	67	46	7	−14	−4	8
relative to the actual variance	81%	55%	8%	−17%	−5%	9%
Correlation coefficient between XS and estimate (ideal value: 1)	0.69	0.78	0.92	0.89	0.92	0.95
Standard-deviation of the differences (ideal value: 0)	7.0	5.8	3.6	4.5	3.7	2.7
relative to the mean of XS value	13%	11%	6%	8%	7%	5%

(likely due to not taking into account calibration coefficients), and the correlation coefficient is low, especially in band XS3. The PCA method introduces too much structure from the P band into the XS1 and XS2 bands. The bias is small and the best results are obtained for XS3 band. The results still remain of low quality. The three methods using the ARSIS concept provide similar results which are of good quality and fairly close to the ideal values. ARSIS Model 1 (identity) does not perform so well for the XS3 band because it does not take into account the spectral behavior of the small-size structures which are set up equal to those of the P band. ARSIS Model 2 and ARSIS-RWM perform better. Tables 2 through 4 prove the large potential of the ARSIS concept in the synthesis of multispectral images.

Table 5 shows the performance of each method in synthesizing the most frequent actual triplets. The total of pixels they represent amounts to 23 percent of the total number of pixels in the image. Hence, synthesizing them accurately is of primary importance for classification purposes. In this Table, for each of these triplets, the number of pixels carrying this triplet in the synthesized images is compared to the corresponding number in the original images. The differences are summed up for all the triplets, giving the "difference with original" in Table 5. A difference equal to 0 means that the geographical location of the triplets is exactly the same as in the original images. Because of the bias, the Brovey transform is unable to retrieve any of these triplets. Very poor results were also obtained with the IHS method: it retrieves only 721 of the 1,675 triplets (43 percent) and only 12 percent of the corresponding pixels. This means that it does not synthesize correctly the triplets, but even those it does retrieve are not correctly allotted to the pixels: this would induce errors in cartography after classification. The PCA method provides fairly good results: 14 percent of the pixels were not carrying the correct triplets. The best results, with close to ideal values, were attained by the methods within the ARSIS concept: all the triplets were exactly retrieved and the number of retrieved pixels carrying one of those triplets in both the original and synthesized images was almost exactly the same. This ensures on the one hand a good classification, and on the other hand a good accuracy in mapping from this classification.

Munechika *et al.* (1993) and Wald *et al.* (1997) computed a total error to summarize some of the various Tables. We generalized their formulas in order to be able to compare errors obtained from the different methods, different cases, and different sensors. Let M be the mean radiance of the N original spectral images B_i then,

$$M = (1/N) \sum_{i=1}^{N} B_i.$$

TABLE 5. PERFORMANCE IN SYNTHESIZING THE MULTISPECTRAL INFORMATION. DIFFERENCE BETWEEN THE ACTUAL FREQUENCY OF A TRIPLET (XS1, XS2, XS3) AND ITS ESTIMATE. ONLY THE MOST FREQUENT TRIPLETS ARE TAKEN INTO ACCOUNT. EACH TRIPLET HAS A FREQUENCY OF AT LEAST 26 PIXELS (0.01 PERCENT). THE TOTAL OF PIXELS THEY REPRESENT AMOUNTS TO 23 PERCENT OF THE TOTAL NUMBER OF PIXELS IN THE IMAGE.

	original	Brovey	IHS	PCA	ARSIS Model 1	ARSIS Model 2	ARSIS RWM
number of predominant triplets	1675	0	721	1673	1675	1675	1675
difference with original (ideal: 0)	—	1675	954	2	0	0	0
(in %)		100%	57%	0%	0%	0%	0%
number of pixels	60372	0	6961	52186	53876	60002	60195
difference with original (ideal: 0)	—	60372	53411	8186	1996	370	177
(in %)		100%	88%	14%	3%	1%	0%

The root-mean-square error RMSE(B_i) for each spectral band can be computed from Tables 2 to 4 as

$$\text{RMSE}(B_i)^2 = \text{bias}(B_i)^2 + \text{standard deviation}(B_i)^2.$$

The relative average spectral error RASE is expressed in percent and characterizes the average performance of a method in the considered spectral bands: i.e.,

$$\text{RASE} = \frac{100}{\text{M}} \sqrt{\frac{1}{N} \sum_{i=1}^{N} \text{RMSE}\ (B_i)}.$$

Table 6 presents the relative average spectral error for each of the methods presented in this paper, plus those discussed by Wald *et al.* (1997), that is, the duplication method and the P + XS method from CNES. The lower the error, the better the method. This RASE parameter allows a comparison at a glance of all methods and of others which are not presented here. For example, Blanc *et al.* (1998b) merged Landsat TM bands 1, 2, and 5 with SPOT-P using iterated filter banks and compared the results to a bi-cubic interpolation of the Landsat TM images. They were not satisfied with the filter banks results nor with the interpolation. From their work, we deduced an RASE of 13.5 for the iterated filter banks and 15.3 for the interpolation. This quantifies the dissatisfaction of these authors, although comparisons should be made with care because the ratio of the spatial resolutions is 1/3 for Landsat TM and P, instead of 1/2 for SPOT-XS and P, and this would negatively impact on the RASE.

The methods are ranked in Table 6 according to the RASE. It perfectly summarizes the previous conclusions, which the authors found valid for several other cases. These conclusions are supported by other authors: Chavez *et al.* (1991), Mangolini *et al.* (1993), Munechika *et al.* (1993), Terretaz (1997), Wald *et al.* (1997), Blanc *et al.* (1998b), Raptis *et al.* (1998), and Zhou *et al.* (1998). The Brovey transform is not relevant at all, mostly because there is a strong bias error due to its very construction. Though it can be partly corrected, it will never reproduce the spectral content in an accurate way, except in rare cases. The IHS method often produces nice looking results but not always as in this case, and it strongly distorts the spectral content of the synthesized images. Wiemker *et al.* (1998) found that the IHS results are inferior to results obtained from the relative spectral contribution method. The same conclusions hold for the PCA method which is, however, far better and can be applied in a more general fashion when compared to the IHS method. Accordingly, the PCA method should be recommended instead of the IHS method. Surprisingly, the duplication method provides fairly good results although it does not call at all on the high-resolution image; Wald *et al.* (1997) underlined that this method may be preferred to the P + XS method on a case by case basis, e.g., if the synthesis of the most predominant triplets and associated pixels is central in the application. The P + XS method from CNES, which is limited to the SPOT case, comes next. Though the authors did not perform an intensive comparison, they believe that interpolation methods, e.g., bi-cubic interpolation, provide a better RASE than does the P + XS method, although the spatial information of the image of higher spatial resolution is not used. However, the effective visual enhancement performed by the P + XS method should be recognized. Finally, the methods calling upon the ARSIS concept provide the best results; Terretaz (1997) underlined the good results obtained using the HPF method, which is one of the possible implementations of the ARSIS concept. The three models (Models 1 and 2 and RWM) give very accurate results (see Tables 2 to 5) and demonstrate the potential of the ARSIS concept for the fusion of images. All published comparisons with existing methods show that the ARSIS concept leads to the best presently achievable results.

TABLE 6. RELATIVE AVERAGE SPECTRAL ERROR (RASE) FOR EACH OF THE METHODS (IN %). VALUES FOR "DUPLICATION," "CNES" (P+XS METHOD), AND ARSIS-RWM ARE DEDUCED FROM WALD *ET AL.* (1997).

Brovey	IHS	PCA	Duplication	CNES	ARSIS Model 1	ARSIS Model 2	ARSIS RWM
65.5	12.8	8.6	7.6	6.7	6.1	5.2	4.1

The implementation of the ARSIS concept presented in this article has been applied to several other published cases in which the synthesis of the spectral content was of importance and synthesized images were further processed:

- SPOT1-3 P and XS images (Mangolini *et al.* 1992; Garguet-Duport *et al.*, 1994; Garguet-Duport *et al.*, 1996);
- SPOT4 P and XS images (Ranchin and Wald, 1996);
- simulated SPOT5 P and multispectral B*i* images (Couloigner *et al.*, 1998a; Couloigner *et al.*, 1998b), or with airborne images (Ranchin and Wald, 1998);
- Landsat TM6 (120 m) and other TM bands (30 m) (Mangolini *et al.*, 1992; Ranchin, 1993; Wald and Baleynaud, 1999); and
- Landsat TM and SPOT P images (Blanc *et al.*, 1998b; Yocky, 1996; Terretaz, 1997; Raptis *et al.*, 1998; Zhou *et al.*, 1998).

Application of the ARSIS Concept to KVR and SPOT Images

In the previous examples, the ratio between h and l is 2 (SPOT) or 4 (Landsat TM). Of course, the ARSIS concept can be applied to other ratios. In this section, we present an application of this concept to XS and KVR-1000 images for which the ratio is 10. Plate 1 displays a color composite of a SPOT XS sub-scene of the city of Riyadh (Saudi Arabia) acquired on 16 May 1993. In this picture, one can clearly see the large interchange of two urban highways, as well as the lots in the sandy areas (upper part) ready for further constructions of buildings. The large white-blue rectangular shape in the lower part is a mall. Figure 10 exhibits a panchromatic image (0.51 to 0.71 μm) taken by the Russian KVR-1000 camera on 07 September 1992. This image has a spatial resolution of 2 m and was acquired more than eight months before the SPOT scene. It displays much more detail

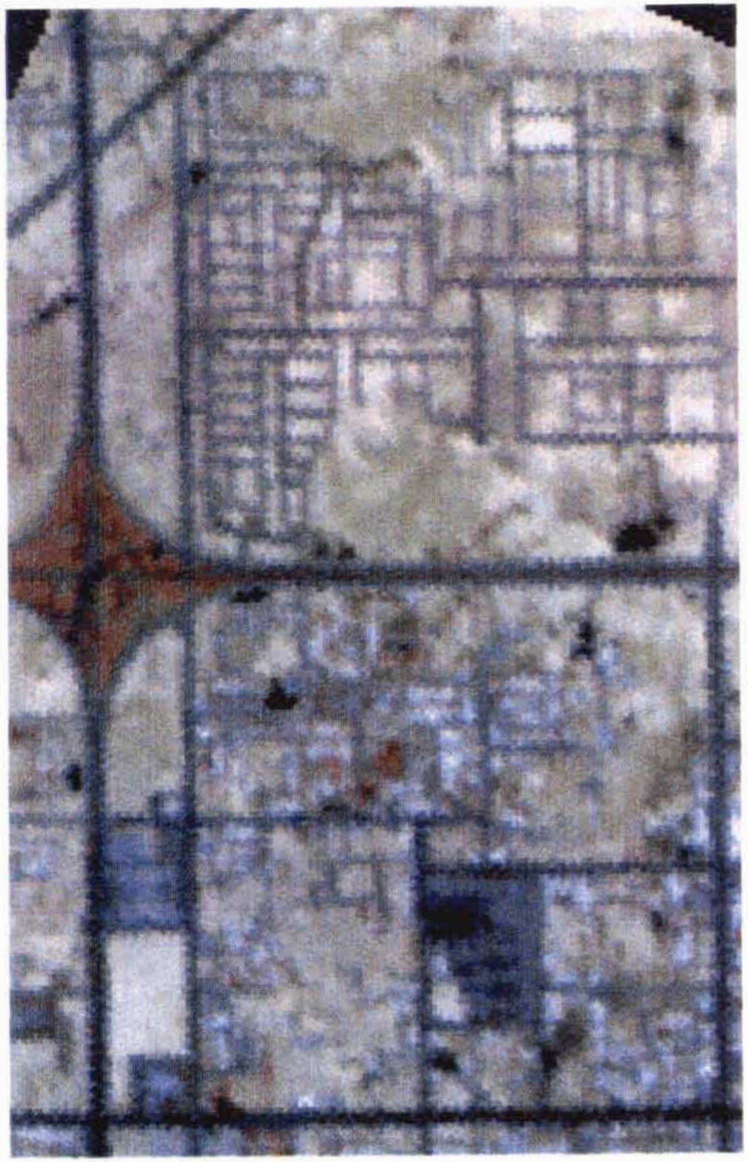

Plate 1 Color composite of a SPOT XS sub-scene of the city of Riyadh (Saudi Arabia) acquired on 16 May 1993. Size of image is 4.28 km (N-S) by 2.74 km (E-W). Copyright CNES SPOT-Image 1993.

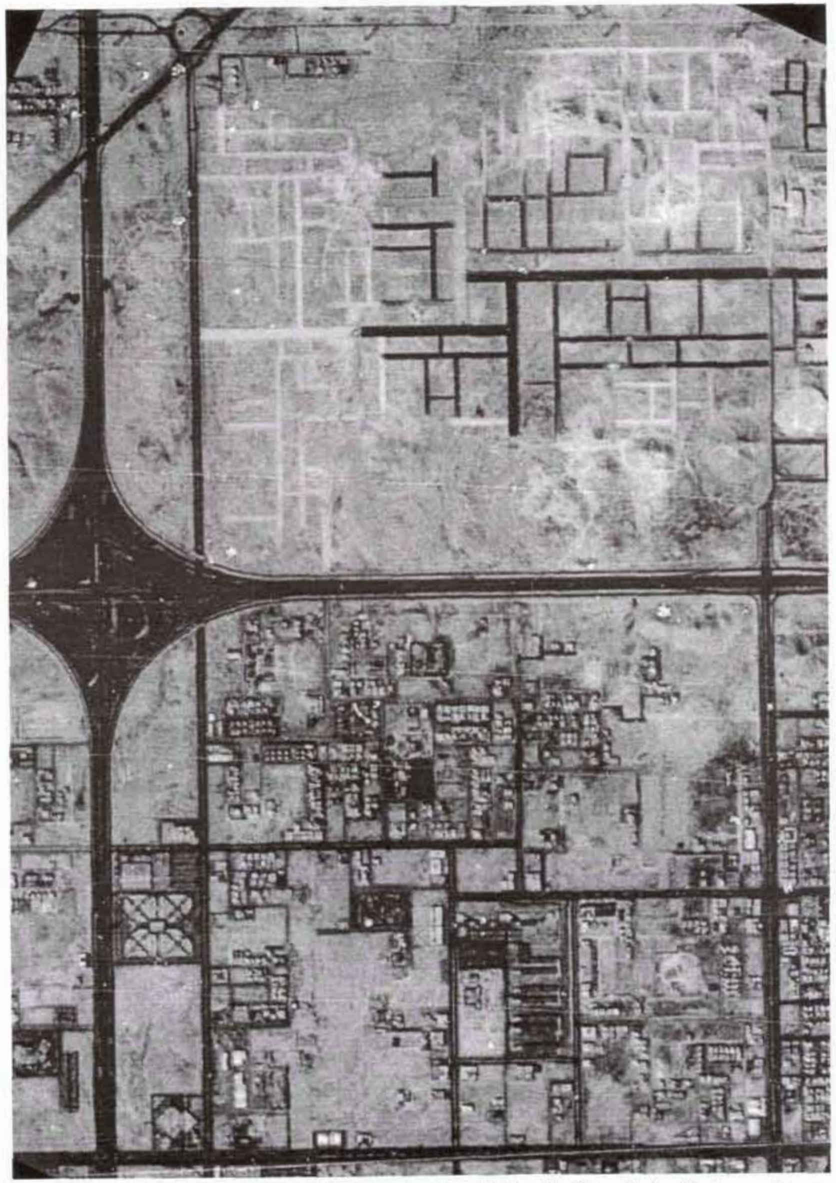

Figure 10. KVR-1000 sub-scene of the city of Riyadh (Saudi Arabia) acquired on 07 September 1992. The area is the same as that in Plate 1.

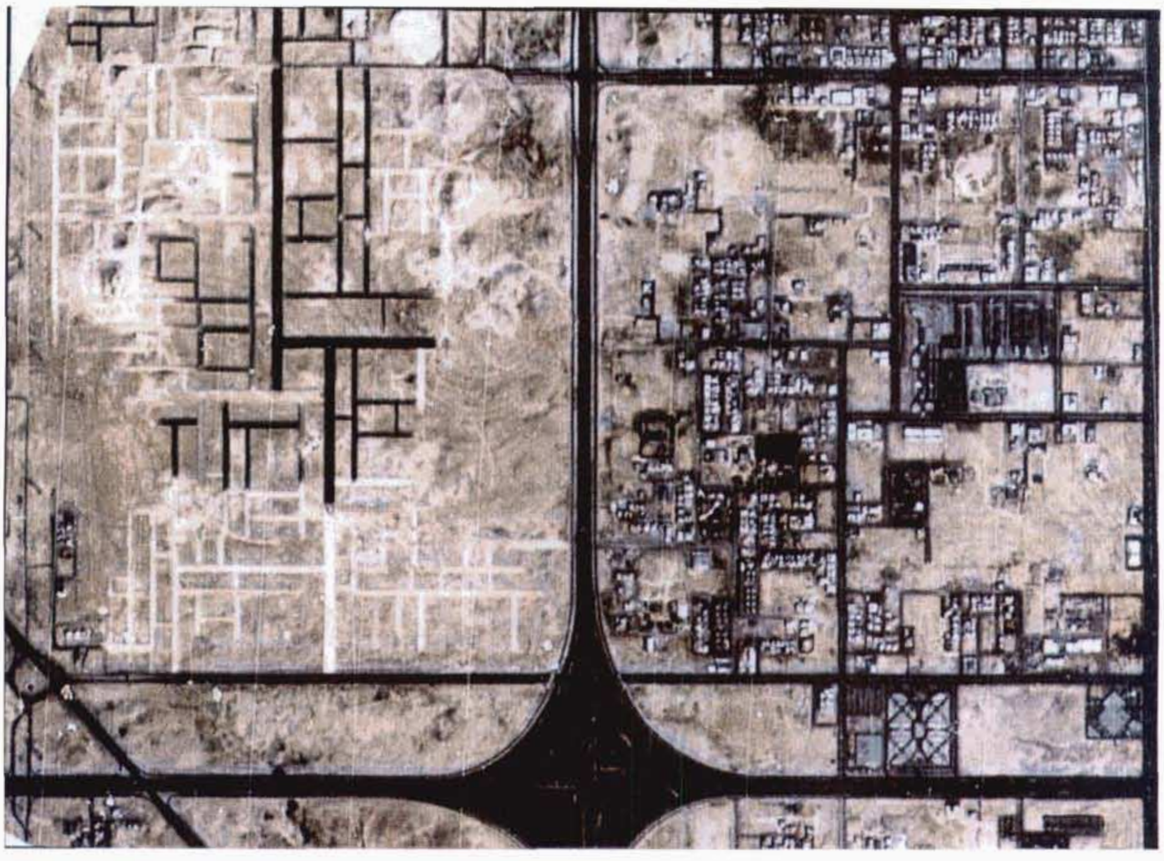

Plate 2 Color composite of the three images synthesized using the IHS method applied to the KVR-1000 and the XS images.

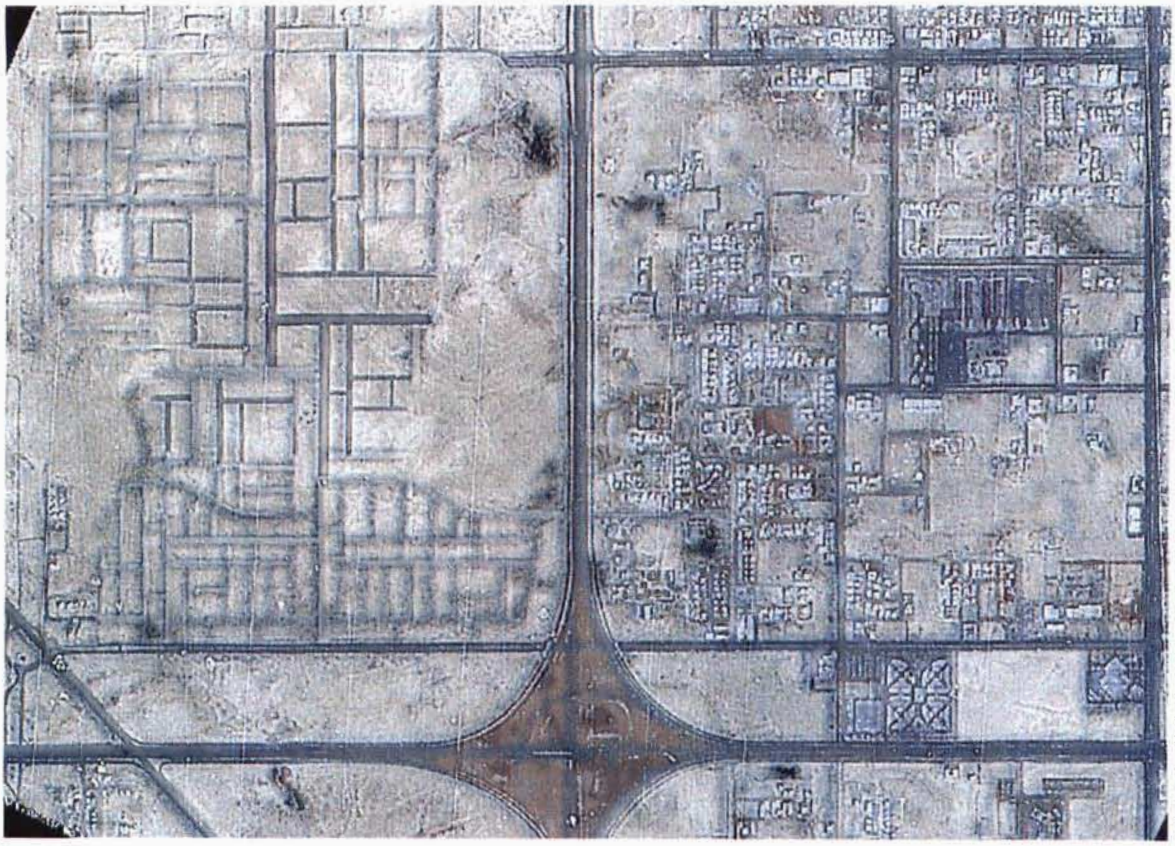

Plate 3 Color composition of the synthesized XS images at 2 m, using the ARSIS concept.

than the SPOT XS image. But there are also two striking features. First, the details of the highway interchange do not appear in the KVR image, likely because of some saturation and defects in the KVR film which had been digitized. Second, the lots in the sandy areas were not yet there at that date: they are not visible in the KVR image. Plate 2 displays the color composite of the three images synthesized using the IHS method applied to the KVR-1000 and the XS images. More details appear compared to the original XS color composite (see, e.g., the mall). However, the IHS procedure has introduced two large defects: the highway interchange is less visible (because of the defects in KVR) and the lots have disappeared (because of the time lag between both images).

The application of the ARSIS concept allows the computation of XS images at the spatial resolution of 2 m. Plate 3 shows a color composite of the three synthesized images. Due to the small details which appear in the synthesized image, it is possible to distinguish the structure of the mall and all of the buildings in this area. The high-quality transformation of the spectral content of XS images, when increasing the resolution, allows the application of a classifier, automatic or not, in order to extract the roads and the buildings. Hence, these synthesized images can be used for classification, or for other methods that need to use the multispectral content provided by the whole set of images with the best spatial resolution available. Ranchin and Wald (1997b) have shown the improvement brought by the use of the ARSIS method to extract roads in urban areas by means of classification methods. One can see that Plate 3 is closer to the XS original composition than the image resulting from the IHS method. Even if the geometrical quality of both images seems to be very close, the images resulting from the application of the ARSIS concept allow one to see all the roads on the interchange, including the details of the lower left loop. The lots which are not visible in the KVR image and in the resulting IHS image are fully visible in the case of the application of the ARSIS concept as they are in the original XS color composition.

Conclusion

In this paper, the ARSIS concept was presented. It is a general framework for the improvement of the spatial resolution of multispectral images. The application of this concept leads to the construction of high spatial resolution multispectral images which are close to the images that the corresponding sensor with the highest resolution would observe. Different methods can be developed based on this concept, depending upon the transformation model, the wavelet transform, and the multiresolution analysis algorithm. The quality of the synthesized images was demonstrated to be the best achievable compared to other fusion methods currently available, in accordance with other authors (Terretaz, 1997; Wald *et al.*, 1997; Raptis *et al.*, 1998; Zhou *et al.*, 1998). The modeling of the spectral behavior of the small-size structures is central to the ARSIS concept. The models presently available are rather straightforward. Although they already produce satisfactory results, better than other methods, efforts should be made to improve them and, finally, provide better synthesized images.

The ARSIS concept is now well understood and is now employed in applications such as urban mapping (Terretaz, 1997; Raptis *et al.*, 1998; Couloigner *et al.*, 1998b), air quality in cities (Wald and Baleynaud, 1999), or detection of equilibrium lines in glaciers (Seidel and Rehauer, personal communication). It is a good and open framework with still many opportunities for the development of different applications and approaches for implementation. This concept will also be applicable to future satellite missions, and, more particularly, to the merging of IRS 1C Pan and LISS images, to the merging of IKONOS panchromatic and multispectral images, and to the use of the resulting images in new applications. For these missions to come, the ratio of spatial resolution between the multispectral set of images and the panchromatic image is 2 or 4, as in the cases of SPOT P and XS or Landsat TM6 and Landsat TM, and is well-suited for the application of the implementation of the ARSIS concept presented in this article.

Acknowledgments

This study partly synthesizes several works performed under the auspices of Aérospatiale, the French Ministry of Defense, CNES, and SPOT-Image.

References

Anonymous, 1986. *Guide des utilisateurs de données SPOT*, three volumes, Editeurs CNES et SPOT-Image, Toulouse, France.

Blanc, Ph., L. Wald, and T. Ranchin, 1998a. Importance and effect of co-registration quality in an example of "pixel to pixel" fusion process, *Proceedings of the Second Conference, "Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images."* (T. Ranchin and L. Wald, editors), Sophia Antipolis, France, 28–30 January, published by SEE/URISCA, Nice, France, pp. 67–74.

Blanc, Ph., T. Blu, T. Ranchin, L. Wald, and R. Aloisi, 1998b. Using iterated rational filter banks within the ARSIS concept for producing 10 m Landsat multispectral images, *International Journal of Remote Sensing*, 19(12):2,331–2,343.

Carper, W.J., T.M. Lillesand, and R. W. Kiefer, 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., S.C. Sides, and J.A. Anderson, 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.

Couloigner, I., T. Ranchin, and L. Wald, 1998a. Benefit of data fusion to urban roads mapping, *Proceedings of the Second Conference "Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images,"* (T. Ranchin and L. Wald, editors), Sophia Antipolis, France, 28–30 January, published by SEE/URISCA, Nice, France, pp. 183–190.

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

Daubechies, I., 1988. Orthonormal bases of compactly supported wavelets, *Communications on Pure and Applied Mathematics*, 41:909–906.

———, 1992. *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics 61, SIAM, Philadelphia, 357 p.

Garguet-Duport, B., J. Girel, and G. Pautou, 1994. Analyse spatiale d'une zone alluviale par une nouvelle méthode de fusion d'images SPOT multispectrales (XS) et SPOT panchromatique (P), *Compte-Rendus de l'Académie des Sciences de Paris*, Sciences de la Vie, 317:194–201.

Garguet-Duport, B., J. Girel, J.-M. Chassery, and G. Pautou, 1996. The use of multiresolution analysis and wavelets transform for merging SPOT panchromatic and multispectral image data, *Photogrammetric Engineering & Remote Sensing*, 62(9):1,057–1,066.

Guo, L.J., and J.McM. Moore, 1998. Pixel block intensity modulation: adding spatial detail to TM band 6 thermal imagery, *International Journal of Remote Sensing*, 19(13):2,477–2,491.

Iverson, A.E., and J.R. Lersch, 1994. Adaptive image sharpening using multiresolution representations. *Proceedings of SPIE Int. Symposium Optical Engineering in Aerospace Sensing*, 2231:72–83.

Li, H., B.S. Manjunath, and S.K. Mitra, 1995. Multisensor image fusion using the wavelet transform, *Graphical Models and Image Processing*, 57(3):235–245.

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.

Mangolini, M., T. Ranchin, and L. Wald, 1992. *Procédé et dispositif pour l'amélioration de la résolution spatiale d'images à partir d'autres images de meilleure résolution spatiale*, French Patent No 92-13961, 20 November 1992.

———, 1993. Fusion d'images SPOT multispectrales (XS) et panchromatique (P), et d'images radar, *De l'optique au radar, les applications de SPOT et ERS*, Cépaduès-Editions, 111 rue Vauquelin, Toulouse, France, pp. 199–209.

Meyer, Y., 1990. *Ondelettes et opérateurs 1: Ondelettes*, Hermann, Paris, France, 215 p.

Munechika, C.K., J.S. Warnick, C. Salvaggio, and J.R. Schott, 1993. Resolution enhancement of multispectral image data to improve classification accuracy, *Photogrammetric Engineering & Remote Sensing*, 59(1):67–72.

Pohl, C., and J.L. van Genderen, 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, *Earth Remote Sensing Using the Landsat Thematic Mapper and SPOT Systems, Proceedings SPIE Conference*, 660:78–102.

Ranchin, T., 1993. *Applications de la transformée en ondelettes et de l'analyse multirésolution au traitement des images de télédétection*, Thèse de Doctorat en Sciences de l'Ingénieur, Université de Nice, Sophia Antipolis, 146 p.

Ranchin, T., and L. Wald, 1993. The wavelet transform for the analysis of remotely sensed images, *International Journal of Remote Sensing*, 14(3):615–619.

———, 1996. Preparation of SPOT 4 mission: Production of high resolution (10 m) multispectral images using the ARSIS method, *Proceedings of 15th Symposium of EARSeL, Progress in Environmental Research and Applications* (E. Parlow, editor), Basel, Switzerland, 04–06 September 1995, A.A. Balkema, Rotterdam, Brookfield, pp. 175–179.

———, 1997a. Data fusion of remotely sensed images using the wavelet transform: The ARSIS solution, *Proceedings of SPIE 1997 International Symposium on Optics, Imaging and Instrumentation. Mathematical Imaging: Wavelet Applications in Signal and Image Processing V*, San Diego, California, 28 July–02 August pp. 272–280.

———, 1997b. Fusion d'images HRV de SPOT panchromatique et multibande à l'aide de la méthode ARSIS: apports à la cartographie urbaine, *Actes des sixièmes journées scientifiques du réseau Télédétection de l'AUPELF-UREF; Télédétection des milieux urbains et périurbains*, Liège, Belgium, 02–05 October 1995, pp. 283–290.

———, 1998. Fusion of airborne and spaceborne images in visible range, *Proceedings of 18th Symposium of EARSeL, Operational Remote Sensing for Sustainable Development*, Enschede, The Netherlands, 11–14 May, in press.

Ranchin, T., L. Wald, and M. Mangolini, 1994. Efficient data fusion using wavelet transforms: The case of SPOT satellite images, *Proceedings of SPIE 1993 International Symposium on Optics, Imaging and Instrumentation. Mathematical Imaging: Wavelet Applications in Signal and Image Processing*, San Diego, California, United States of America, 11–16 July 1993, 2034:171–178.

———, 1996. The ARSIS method: A general solution for improving spatial resolution of images by the means of sensor fusion, *Proceedings of the First Conference Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images*, (T. Ranchin and L. Wald, editors), Cannes, France, 06–08 February 1996, published by SEE/URISCA, Nice, France, pp. 53–58.

Raptis, V.S., R.A. Vaughan, T. Ranchin, and L. Wald, 1998. Assessment of different data fusion methods for the classification of an urban environment, *Proceedings of the Second Conference, Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images*, (T. Ranchin and L. Wald, editors), Sophia Antipolis, France, 28–30 January 1998, published by SEE/URISCA, Nice, France, pp. 167–182.

Terretaz, P., 1997. Comparison of different methods to merge SPOT P and XS data: Evaluation in an urban area, *Proceedings of 17th Symposium of EARSeL, Future Trends in Remote Sensing*, (P. Gudmansen, editor), Lyngby, Denmark, 17–20 June 1997, A.A. Balkema, Rotterdam, Brookfield, pp. 435–445.

Tom, V.T., 1987. *System for and Method of Enhancing Images Using a Multiband Information*, USA Patent 4,683,496, 28 July.

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

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

Wiemker, R., B. Prinz, G. Meister, R. Franck, and H. Spitzer, 1998. Accuracy assessment of vegetation monitoring with high spatial resolution satellite imagery, *International Archives of Photogrammetry and Remote Sensing*, 32 (Part 7):285–292.

Yocky, D.A., 1996. Multiresolution wavelet decomposition image merger of Landsat Thematic Mapper and SPOT panchromatic data, *Photogrammetric Engineering & Remote Sensing*, 62(9):1067–1074.

Zhou, J., D.L. Civco, and J.A. Silander, 1998. A wavelet method to merge Landsat TM and SPOT panchromatic data, *International Journal of Remote Sensing*, 19(4):743–757.

(Received 06 July 1998; revised and accepted 14 February 1999)