

Synthesis of multispectral images to high spatial resolution: a critical review of fusion methods based on remote sensing physics

Claire Thomas, Thierry Ranchin, Lucien Wald, Jocelyn Chanussot

▶ **To cite this version:**

Claire Thomas, Thierry Ranchin, Lucien Wald, Jocelyn Chanussot. Synthesis of multispectral images to high spatial resolution: a critical review of fusion methods based on remote sensing physics. IEEE Transactions on Geoscience and Remote Sensing, 2008, 46 (5), pp.1301-1312. 10.1109/TGRS.2007.912448 . hal-00348848

HAL Id: hal-00348848
https://hal.science/hal-00348848

Submitted on 22 Dec 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Synthesis of Multispectral Images to High Spatial Resolution: A Critical Review of Fusion Methods Based on Remote Sensing Physics

Claire Thomas, Thierry Ranchin, *Member, IEEE*, Lucien Wald, and Jocelyn Chanussot, *Senior Member, IEEE*

***Abstract*—Our framework is the synthesis of multispectral images (MS) at higher spatial resolution, which should be as close as possible to those that would have been acquired by the corresponding sensors if they had this high resolution. This synthesis is performed with the help of a high spatial but low spectral resolution image: the panchromatic (Pan) image. The fusion of the Pan and MS images is classically referred as pan-sharpening. A fused product reaches good quality only if the characteristics and differences between input images are taken into account. Dissimilarities existing between these two data sets originate from two causes—different times and different spectral bands of acquisition. Remote sensing physics should be carefully considered while designing the fusion process. Because of the complexity of physics and the large number of unknowns, authors are led to make assumptions to drive their development. Weaknesses and strengths of each reported method are raised and confronted to these physical constraints. The conclusion of this critical survey of literature is that the choice in the assumptions for the development of a method is crucial, with the risk to drastically weaken fusion performance. It is also shown that the Amélioration de la Résolution Spatiale par Injection de Structures concept prevents from introducing spectral distortion into fused products and offers a reliable framework for further developments.**

***Index Terms*—Image enhancement, image processing, merging, multiresolution techniques, remote sensing.**

I. Introduction

ALL IMAGING applications that require analysis of two or more images of a scene can benefit from image fusion. We rely on two definitions extracted from literature—Wald's [1] and Piella's [2] definitions, respectively. Wald [1] defines image fusion as "a formal framework in which are expressed means and tools for the alliance of data originating from different sources. It aims at obtaining information of a greater quality, although the exact definition of 'greater quality' will depend on the application". According to Piella [2], fusion is "the combination of pertinent (or salient) information in order to synthesize an image more informative and more suitable for visual perception or computer processing," where the "pertinence" of the information is also dependent on the application task.

Manuscript received May 29, 2007; revised July 31, 2007. This work was supported in part by the French Ministry of Defense under Grant 0534045.

C. Thomas, T. Ranchin, and L. Wald are with the Ecole des Mines de Paris/Armines, 06904 Sophia Antipolis, France (e-mail: Lucien.wald@ensmp.fr).

J. Chanussot is with the Grenoble Images Speech Signals and Automatics Laboratory, Grenoble Institute of Technology, 38031 Grenoble, France (e-mail: jocelyn.chanussot@lis.inpg.fr).

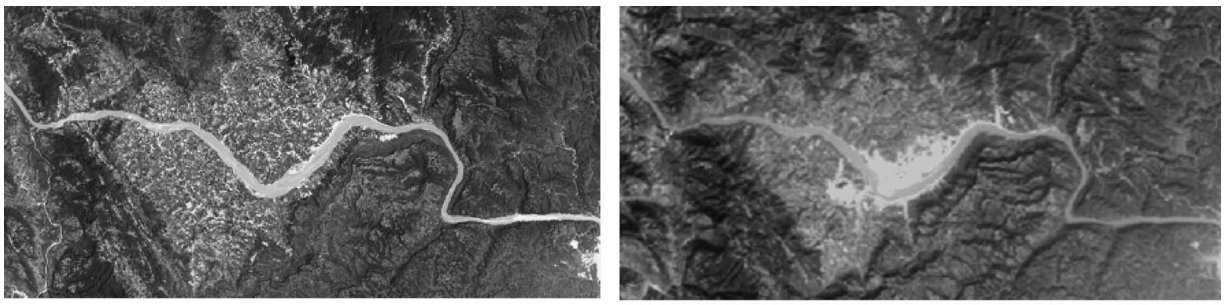
Digital Object Identifier 10.1109/TGRS.2007.912448

Each application field of image fusion leads to an interpretation of these definitions and also involves specific physical considerations. In this paper, we focus on a particular application field of image fusion in remote sensing, which is the synthesis of multispectral (MS) images to the higher spatial resolution of the panchromatic (Pan) image. The main spectral characteristic of the Pan modality is to cover a broad range of the wavelength spectrum, whereas an MS band covers only a narrow spectral range. Since more energy comes to Pan sensor, time acquisition can be reduced still preserving the same intensity response as MS images in terms of the number of photons. The advantage of the Pan image is a smaller size of pixels and, hence, better spatial resolution. The Pan image, thus, combines low spectral resolution and high spatial resolution, whereas the MS image combines reverse characteristics. The design of MS sensors with better resolution is limited by technical constraints of onboard storage and bandwidth transmission of the images from the satellite to the ground.

Therefore, due to a combination of observational constraints imposed by the acquisition system, spaceborne imagery usually provides separated but complementary product types. An increasing number of applications such as feature detection [3] or land cover classification [4] require high spatial and high spectral resolution at the same time for improved classification results, strengthened reliability, and/or a better visual interpretation. In response to those requirements, image fusion has become a powerful solution to provide an image containing the spectral content of the original MS images with enhanced spatial resolution. This particular field of application of data fusion is usually called pan-sharpening. More precisely, the framework of the presented study is the synthesis of fused MS images that should be as close as possible to those that would have been observed if the corresponding sensors had this high spatial resolution.

Even with geometrically registered Pan and MS images, dissimilarities might exist between these modalities. In addition to changes produced by their different spectral acquisition bands of Pan and MS images, drastic changes might also occur in the scene for two different acquisition times. Many authors attempt to figure out relationships between these remotely sensed images for the development of their fusion method. However, because of variations between these images, no obvious universal link exists. This is the emphasis of this paper, which demonstrates that physics must be taken into account and discusses ways to do so. The reliability of the starting assumption adopted by several publications is discussed and

0196-2892/$25.00 © 2008 IEEE



(a) (b)

Fig. 1. SPOT2, Three Gorges dam, China. Illustration of the changes that can be observed between two acquisition times. (a) Pan modality, 10 m, acquired in 1990. (b) Green modality $XS1$, 20 m, 1998. Copyright CNES SPOT-Image 1990 and 1998.

confronted to physics. The purpose of this critical survey is to highlight the domain of validity and the shortcomings of such approaches compared to others, leading to recommendations for adopting existing methods or developing new ones.

This paper is organized as follows. Section II illustrates the effects of the environmental physics in MS and Pan images to highlight the complexity of the images for the synthesis of MS images at high resolution using a Pan image. Section III introduces notations, details the challenges of fusion, and explains why authors need to make assumptions to develop their fusion methods. Then, Section IV analyzes the different assumptions of the main categories of fusion methods found in literature to study to what extent physics is taken into account. Section V is dedicated to the most recent developments that are discussed under the light of this analysis. Last, in Section VI, a conclusion is drawn, and a series of recommendations is proposed.

II. Input Image Characteristics

Generally, MS and Pan modalities often display the same geographic area. We assume that Pan and MS input data sets are *a priori* geometrically registered. The task of registration is a very challenging one [5], particularly when images come from different platforms. Blanc *et al.* [6] showed that a geometrical distortion of only 0.1 pixel in standard deviation produces a noticeable effect on the quality of fused images resulting from a pixel-to-pixel fusion process.

Nevertheless, even with perfectly coregistered images, the Pan and MS data sets can locally exhibit some dissimilarities, whose origin is not always well understood by the fusion community [7], [8]. This can impact the quality of fused images. Several types of dissimilarities are illustrated and discussed in the following.

One of the most common dissimilarities is object occultation. The first possible cause of occultation is when the time lag in the acquisition of the Pan and MS images is not close to zero. Several transformations may occur in this lapse of time—variations in vegetation depending on the season, different illumination conditions, development of suburbs, or changes due to natural catastrophes (earthquakes, floods, volcano eruptions, etc.). Fig. 1 displays the effects of a human construction—the Three Gorges dam of the Yangtze River, China. For the past century, the Chinese government has considered ways to harness the strength of the Yangtze River and tame its devastating power. A decade ago, the government set out to accomplish this goal as they began construction of a very large dam to be completed in 2009. The first picture is a SPOT2 Pan modality taken in 1990 with 10 m of spatial resolution. The second one depicts the same scene 8 years later in the green spectral range. The spatial resolution is 20 m. Works began between these two dates, impacting the environment at a large scale. Riverbanks have been cleared from all vegetation to welcome construction works. Changes that occurred in these 8 years are highly perceivable in these two images. Note that we are using the word "occultation"; we may have used as well "disappearance" or inversely "appearance."

Even in the case of almost identical acquisition times, occultation of objects may also occur because of their different spectral bands of acquisition. The demonstration is carried out on Pan images degraded to the original spatial resolution of MS images. The next three illustrations are identically organized. The first column [Fig. 2(a)] depicts the part of a satellite image acquired in the wavelength range of the Pan modality. This part is downsampled to the original spatial resolution of the MS image. The corresponding MS excerpt is presented in Fig. 2(b). Fig. 2(c) displays a particular transect of each image under the form of a graph, delimited in each image thanks to a frame. The curve represents grayscale levels as a function of the pixel number in the extracted line. Pan and blue transects are plotted with full and dashed lines, respectively.

The Pan excerpt of Fig. 2(a) shows a path that is surrounded by vegetation; the image was acquired by Quickbird over the city of Fredericton, Canada. The spatial resolution is 0.7 m, which is downsampled to the original resolution (2.8 m) of MS images. The same portion of the image in blue range is exhibited in Fig. 2(b), with the original resolution of 2.8 m.

The path that is visible in the Pan image is missing in the MS modality. This generates a local dissimilarity between these two images. The reverse case is also possible, with a structure present in the MS modality but which cannot be distinguished in the Pan modality, as displayed in Fig. 3. The landscape of the excerpt shows agricultural fields in the area of Toulouse, France. Images were collected by the satellite SPOT5. Fig. 3(a) is the Pan modality with 2.5 m of spatial resolution and downsampled to 10 m, and Fig. 3(b) depicts the near-infrared modality with resolution of 10 m. Fig. 3(c) exhibits Pan and blue transects.

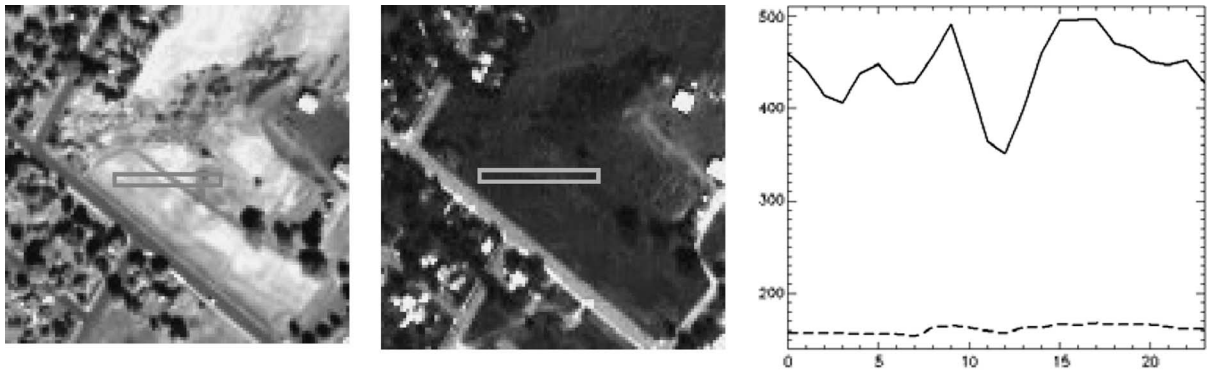


Fig. 2. Quickbird, excerpt of Fredericton, Canada. Illustration of an object occultation in an MS modality. (a) Pan, 0.7 m downsampled to 2.8 m. (b) Blue, 2.8 m. (c) Transects of the segment shown in each image (dashed: blue; full line: Pan). Copyright Digital Globe 2002.

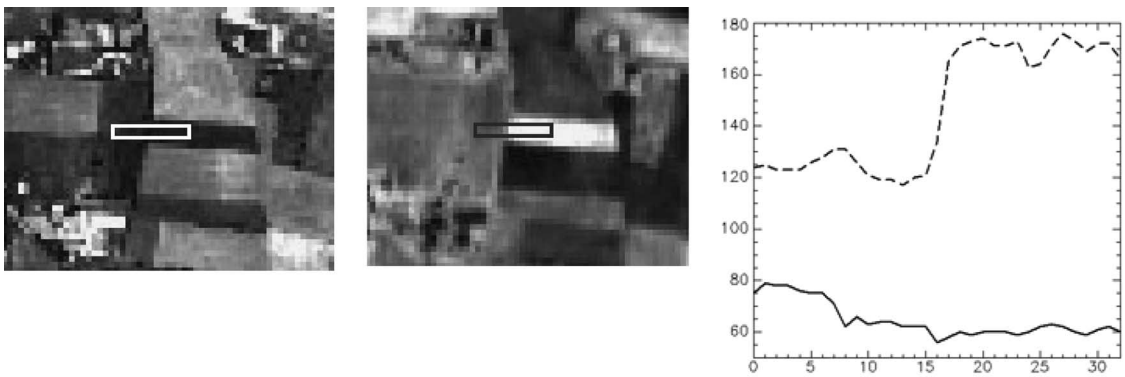


Fig. 3. SPOT5, Toulouse, France. Illustration of an object occultation in the Pan modality. (a) Pan, 2.5 m downsampled to 10 m. (b) Near-infrared, 10 m. (c) Transects of the segment shown in each image (dashed: near-infrared; full line: Pan). Copyright SPOT Image 2002.

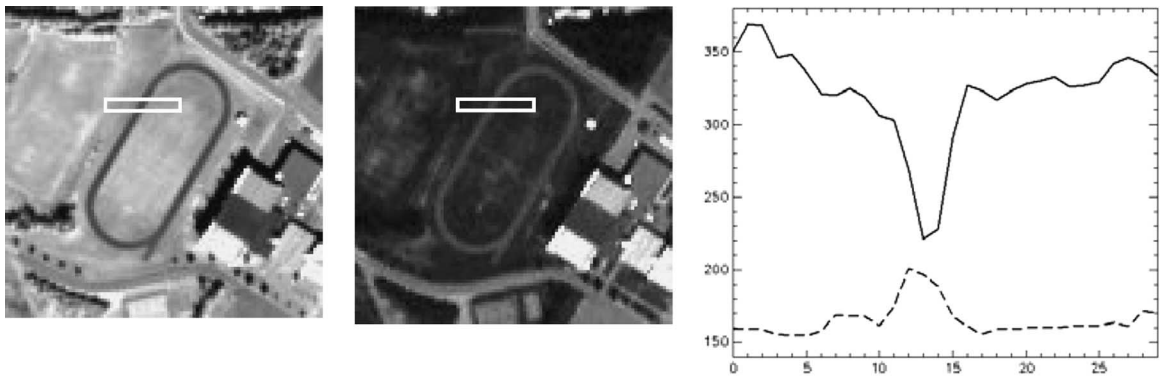


Fig. 4. Quickbird, excerpt of a stadium near downtown Fredericton, Canada. (a) Pan, 0.7 m downsampled to 2.8 m. (b) Blue, 2.8 m. (c) Transects of the segment shown in each image (dashed: blue; full line: Pan). Copyright Digital Globe 2002.

The near-infrared transition lies from low- to high-value grayscales, whereas the transect corresponding to a line extraction of the Pan modality does not exhibit any transition at all.

Fig. 4 highlights another type of dissimilarity called the contrast inversion or the contrast reversal information. The excerpt represents a stadium that is located in downtown Fredericton, Canada, and taken by the satellite Quickbird. Fig. 4(a) is the Pan modality at 0.7 m downsampled to 2.8 m, and b is blue with 2.8-m original spatial resolution.

Transects of Fig. 4(c) show that the athletic trail and its surrounding grass are in contrast inversion in the two excerpts (darker or brighter feature, respectively).

Even if the Pan and MS images are said to be acquired at the same time, acquisition moments are not precisely identical. Moreover, the two sensors do not exactly aim at the same direction. These two facts have an impact on the imaging of fast-moving objects. For example, the position of a vehicle might slightly differ from one image to another. Geometrically

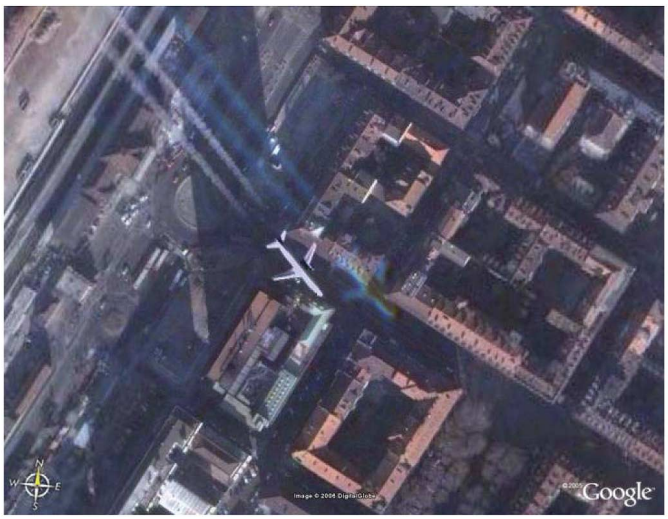


Fig. 5. True-color composition of fused Quickbird images (0.7 m) extracted from Google Earth over the city of Turin, Italy. Illustration of the default of ghost caused by fast-moving objects. Copyright Google.

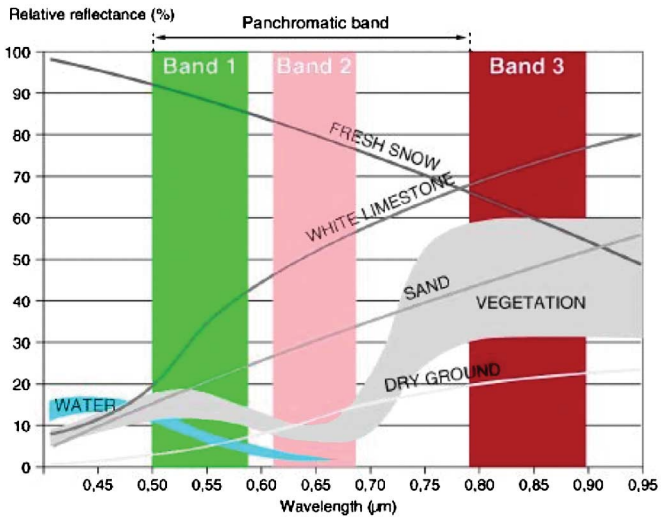


Fig. 6. Sketch of the normalized spectral responses of the different spectral bands of SPOT 1-2-3 and relative reflectance of several ground types as a function of the wavelength (in micrometers).

registered images superimpose most structures located in the whole images. Moving objects such as cars, lorries, or planes are locally no longer superimposable. The consequence on fusion is that low-resolution moving objects might create a ghost corresponding to the low-resolution version of the moving object. Fig. 5 comes from Google Earth, a tool that offers a worldwide coverage of fused products. The picture shows a plane over the city of Turin, Italy.

This particular excerpt is a Quickbird fused product at the spatial resolution of the Pan modality (0.7 m). The high-resolution plane is in bright white; however, a low-resolution multicolor and slightly shifted ghost of this plane also appeared during the fusion process.

Many physical phenomena interfere during the acquisition phase. They are not always well understood by the fusion community. As already discussed, for images acquired at two different acquisition times, variation of illumination in the scene may be observed when hours are different and when the sun declination is different depending on the season. It causes a modification in the size and orientation of shadows, and changes in contrast luminance as well. Shadows produce physical occultation of objects. Changes in ground occupation may also occur, and objects may move (Figs. 1 and 5).

In the case of two successive days at the same local time, such as daily data provided by the satellite Formosat-2, shadows and contrasts would be almost unchanged. However, even with clear skies, optical atmospheric conditions may be different. For instance, the load of water vapor or aerosol in the atmosphere column has an impact on the sharpness of the edges, and the corresponding impact is a function of the wavelength.

Different spectral bandwidths cause differences in acquired images (Figs. 2–4). Spectral dissimilarities are the result of the interaction between the incident solar energy and the matter; a sensor collects the resulting energy retransmitted toward the satellite, and, finally, the sensor synthesizes the final image through optical and electronic devices. The matter is actually the soil type (sand, snow, vegetation, bare and dry ground, wet lands, etc.), and the corresponding reflectance varies with the wavelength. Fig. 6 features the spectral response functions of the Pan and the three MS modalities of SPOT 1-2-3 satellites. Green, pink, and red bands correspond to green, red, and near-infrared modalities, respectively. The Pan spectral response of the sensor is not drawn; however, the bandwidth of the Pan modality is indicated on the top of the graph. Spectral reflectances of several ground types are overlaid.

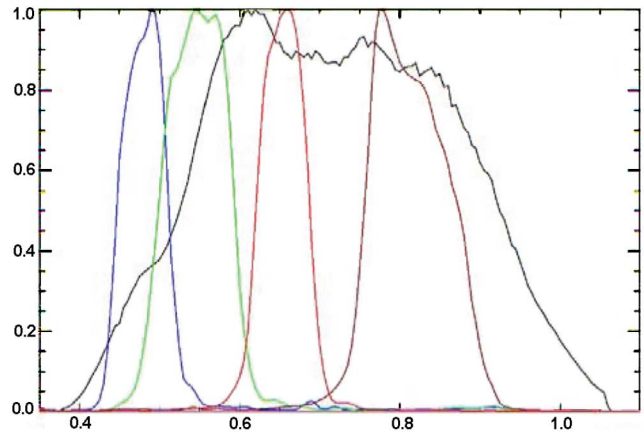


Fig. 7. Normalized spectral responses of the different spectral bands of Ikonos as a function of the wavelength (in micrometers). Blue, green, red, brown, and black curves respectively correspond to blue, green, red, near-infrared, and Pan modalities. Source: http://www.geoeye.com/products/imagery/ikonos/spectral.htm.

This picture shows that the Pan and near-infrared bands do not overlap at all. The Pan spectral range covers a little wider portion of the spectrum toward the red wavelengths. One given pixel gives a nonzero grayscale value if its ground type reflects the incident solar energy in the range of the acquisition bandwidth of the corresponding sensor. In this graph, an object that reflects solar energy in the wavelengths surrounding 0.75 μm is only visible in the Pan modality and completely occulted in MS modalities, as is the case of Fig. 2. On the other hand, if an object exhibits a signal only in the near-infrared range, near-infrared modality can capture this object while Pan cannot, like in Fig. 3.

To explain contrast inversion in Fig. 4, we refer to Ikonos spectral band responses in Fig. 7. This figure shows spectral response functions of blue, green, red, near-infrared, and Pan modalities for the satellite Ikonos corresponding to the colors blue, green, red, brown, and black, respectively.

As the Ikonos Pan spectral band contains a large portion of the near-infrared wavelength range, which is very sensitive to vegetation, areas containing such landscape are brighter than the rest of the scene. On the contrary, vegetation has almost no response in the blue range, which exhibits dark digits in Fig. 4. Consequently, if vegetation is surrounded by another soil type that is characterized by constant response with wavelength, their transition appears in contrast inversion in the two bands.

In conclusion, the MS and Pan data sets might present some local dissimilarities such as object occultation, contrast inversion, or moving objects due to different spectral bands of the sensors or different times of acquisition. These effects are due to environmental physics. If they are not—or only partially—taken into account, fusion success might be endangered by the apparition of artifacts. Before proceeding to a critical analysis of literature on this point, Section III defines some useful notations.

III. Fusion Purpose and Notations

In the following, A defines Pan modality, and B_k is the kth MS modality. N is the total number of MS bands. In the case of Ikonos, where $N = 4$, k equal to 1 designates the blue modality, whereas 4 represents the near-infrared one. The subscript "$_0$" is the Pan image spatial resolution *res0*; thus, the original Pan image is called A_0. The subscript "$_1$" is the resolution index of the original MS images *res1*, giving B_{k1}. "*" designates the fused products: fused MS images at the spatial resolution of Pan *res0* are called $(B_{k1})_0^*$. The issue of the synthesis of MS images to higher spatial resolution can, thus, be mathematically expressed by

$$(B_{k1})_0^* = f(B_{k1}, A_0). \tag{1}$$

The aim of fusion is to perform a high-quality transformation of the MS content when increasing the spatial resolution from *res1* to *res0*. The problem may be seen as the inference of the information that is missing to the images B_{k1} for the construction of the synthesized images $(B_{k1})_0^*$ [9], [10].

The ideal fusion method should be able to correctly handle all dissimilarities described in Section II. Nowadays, no published fusion method actually fulfills the totality of these requirements.

IV. Critical Presentation of the Groups of Fusion Methods

This section is dedicated to the survey of literature concerning the assumptions made by authors when developing their fusion methods. Attention should be paid to the consideration of physics in the fusion process. Our purpose is to demonstrate that several assumptions *a priori* reduce fusion success because they do not consider all the physics of acquisition. We adopt Wald's classification [10] of the fusion methods to make up a brief description of the standard methods. The three categories are projection–substitution, relative spectral contribution, and methods that belong to the Amélioration de la Résolution Spatiale par Injection de Structures (ARSIS) concept. Other classifications of the fusion methods could have been adopted, as discussed later.

A. Projection–Substitution Methods

These fusion methods exploit a vectorial approach since all MS modalities are simultaneously synthesized. The MS modalities are projected into a new space to reduce information redundancy and obtain decorrelated components. One of the components isolates structures of MS images from the rest of the information, mainly related to color. The assumption of this type of methods is that the structures contained in this structural (or geometrical) component are equivalent to those in the Pan modality. Next, the substitution consists of the total or partial replacement of this structural component by the Pan modality. Last, the inverse projection is performed to obtain $(B_{k1})_0^*$.

Most famous projection–substitution methods are based on principal component analysis (PCA) [11] and intensity hue saturation (IHS) [12]. IHS is a transform that originally applies on exactly three MS modalities. In the field of fusion, the space transform is often confused with the fusion method itself. In the following, IHS will, thus, designate the fusion method.

The first step of IHS is the upsampling of the input MS images to adjust to the Pan spatial resolution. These images are noted $(B_{k1})_0^{\text{upsamp}}$ since B_{k1} images are upsampled from *res1* to *res0*. Direct linear IHS transform was originally defined for three modalities

$$(I_1)_0^{\text{upsamp}} = \alpha_1 (B_{1,1})_0^{\text{upsamp}} + \alpha_2 (B_{2,1})_0^{\text{upsamp}} + \alpha_3 (B_{3,1})_0^{\text{upsamp}} \tag{2}$$

where Chavez *et al.* [11] defined $\alpha_i = 1/\sqrt{3}\ \forall i$, and Gonzalez and Woods [13] defined $\alpha_i = 1/3\ \ \forall i$. In this formalism, $(I_1)_0^{\text{upsamp}}$ is a linear combination of the upsampled input MS images, and it is supposed to gather the geometrical structures embedded in these images. The total or partial replacement of $(I_1)_0^{\text{upsamp}}$ requires prior histogram adjustment between the Pan and this component. Let newI be the new intensity. The total replacement of $(I_1)_0^{\text{upsamp}}$ is given by

$$\text{newI} = A_0 \tag{3}$$

and the partial replacement is given by

$$\text{newI} = a(I_1)_0^{\text{upsamp}} + (1-a)A_0, \quad \text{with } a \in [0,1]. \tag{4}$$

Last, reverse transform is applied on unchanged H and S components with newI.

Tu *et al.* [14], [15] pioneered a new way of tackling the linear IHS formulation discussed in [13]. The kth fused MS modality $(B_{k1})_0^*$ is given by

$$(B_{k1})_0^* = (B_{k1})_0^{\text{upsamp}} + d, \quad \text{with } d = \text{newI} - (I_1)_0^{\text{upsamp}}. \tag{5}$$

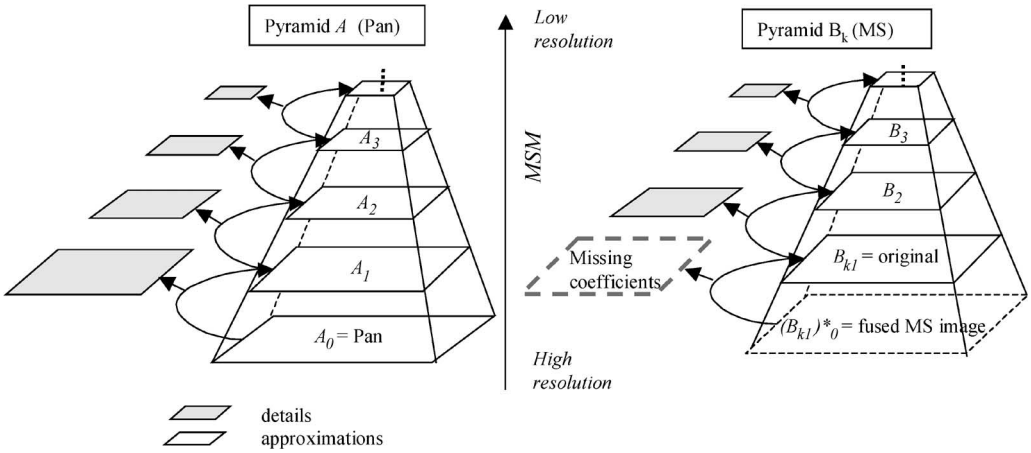


Fig. 8. Hierarchical decomposition of the information (ARSIS concept).

The different implementations of IHS depend on the expression of d. The fused image $(B_{k1})_0^*$ is, thus, a function of newI, i.e., a function of Pan, and of $(I_1)_0^{\text{upsamp}}$ corresponding to a function of all the upsampled low-resolution images.

B. Relative Spectral Contribution Methods

These methods are also based on the linear combination of bands. The fundamental assumption consists of considering that low-resolution Pan can be written as a linear combination of original MS images. This assumption comes from the overlap of the spectral bands

$$A_1 = \Sigma \alpha_k B_{k1}, \quad \text{with } k = \{1, \ldots, N\}. \tag{6}$$

The weights α_k are related to the spectral contribution of the low-resolution version of the Pan modality to each MS band. A filtering operation applied on the Pan image is, thus, implicitly required. Well-known methods are Brovey [16], color normalized, and P+XS [17]. In all these formalisms, the fused MS product is a function of this linear combination and of the Pan image as well.

C. ARSIS Concept Implementations

ARSIS is the French acronym for "Amélioration de la Résolution Spatiale par Injection de Structures" (Improving Spatial Resolution by Structure Injection) [9]. Its fundamental assumption is that the missing spatial information in MS modalities can be inferred from the high frequencies, which lay between *res0* and *res1* of the Pan image, and possibly from external knowledge. Multiscale or multiresolution algorithms are applied onto these images to obtain a scale-by-scale description of the information content of both images, generally represented by a pyramidal structure, as illustrated in Fig. 8. This is called the multiscale model (MSM) [10], [18].

The missing coefficients (dashed plane) to be injected in pyramid B from image A are needed to synthesize the fused image $(B_{k1})_0^*$ located in the missing bottom of pyramid B (dashed plane). Because of the dissimilarities between images, high frequencies extracted from the Pan representation and those of an MS image are not exactly equivalent. In the aim of synthesizing MS images close to those that would have been collected by the corresponding sensor if it had the high spatial resolution, adaptation or transformation should be applied to adjust these details to the MS representation [9], [10]. This is called the intermodality model (IMM) [10] or the interband structure model (IBSM) [18]. Several IMMs have been proposed in [9], [10], [18], and [19]. Some implementations of the concept do not perform any transformation of the high frequencies before injection into MS low-resolution images; this is the case of [20], high-pass filter (HPF) [11], and [21]–[23]. The others estimate the transformation IMM to convert the information provided by the multiscale representation of A into those needed for the synthesis of image $(B_{k1})_0^*$. The estimation is performed at lower resolution where images are simultaneously present in both pyramidal representations. Pan high-frequency structures are then transformed using this relation before insertion in the MS modalities: this is called the high-resolution IMM (HRIMM) or the high-resolution IBSM.

Last, the reverse transform of the multiscale algorithm is applied to synthesize the fused MS image. Such methods are presented in a general way in [9]. The MSM may be based on wavelet transforms [18], [24]–[26], Laplacian pyramid decomposition [18], [20], curvelets [27], contourlets, and local average gradient [28]. The MSM should take into account the modulation transfer functions (MTFs) of the real sensors since the MTFs of the MS sensors may be significantly different from one another and, particularly, are different from the Pan one [29]. MTF-tailored filters would limit frequency discontinuities for the fused product [30].

D. Advantages and Drawbacks of Projection–Substitution and Relative Spectral Contribution Fusion Methods

1) Result Depends on the Correlation: By construction, the higher the correlation between the Pan and each MS modality, the better the success of fusion. If the correlation is high, then an MS modality can be written as an affine function of the Pan modality. Thus, if we refer to the IHS formulation [see (2)], the intensity component also becomes an affine function of the Pan modality. As this component is the one to be replaced by the Pan modality, if an adjustment of histogram is performed between

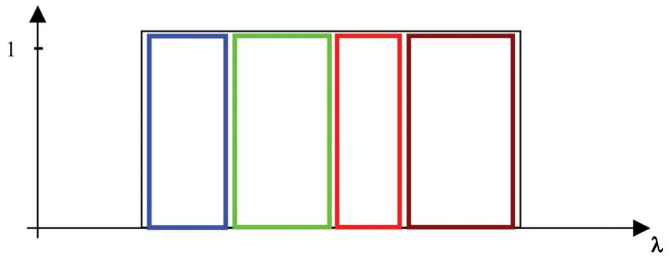


Fig. 9. Ideal normalized spectral responses. In abscissa, the wavelengths are in micrometers.

$(I_1)_0^{\text{upsamp}}$ and Pan, fusion would produce optimal results. The same theoretical proof can be supplied to any linear relationship such as PCA or relative spectral contribution methods. Therefore, the more numerous the dissimilarities between the input images, the worse the quality of the fusion products using these fusion methods.

2) Good Visual/Geometrical Impression in Most Cases: The advantage of these two types of method is to produce a noticeable increase in visual impression with a good geometrical quality regarding the structures [7], [31], [32]. Vijayaraj *et al.* [32] and Yocky [33] stated that they are well suited to certain applications such as cartography or the localization of specific phenomena like target recognition.

3) Spectral Distortion: Their major drawback is spectral distortion, also called the color or radiometric distortion, which is characterized by a trend to present a predominance of a color on the others [24]. This effect is either localized to a certain type of landscape or affecting the whole image. This issue has been also raised by Pellemans *et al.* [34], who postulated that these methods were not adapted to vegetation study. The spectral distortion observed in projection–substitution and relative spectral contribution fused products is due to the modification of the low frequencies of the original MS images [10], [35]. A fused MS modality is mathematically written as a function of the original MS and Pan images. These modalities contain low frequencies, which are injected and alter those of B_{k1}. However, according to the spectral response of recent spaceborne sensors, no obvious relation exists between Pan and MS input modalities, and such a relation is certainly not linear. In Fig. 7, corresponding to Ikonos spectral band responses, if an object reflects solar incident energy in wavelengths located around 1 μm, it will be impossible to infer the pixel value in the Pan modality from other MS modalities since this pixel will have a grayscale value equal to 0 in all MS images. Moreover, this graph shows that blue and green channels overlap, creating spectral redundancy between the two images. The relative spectral contribution assumption would be perfectly true if spectral responses of all the sensors of a satellite verify the ideal theoretical graph from Fig. 9.

As a matter of fact, this assumption is not realistic. Even if the Pan modality was artificially simulated by a combination of actual MS acquisitions, it would never be equal to a linear combination of MS responses. MS airborne or spaceborne sensors do not offer a constant response over the whole bandwidth. Bandwidth limits are characterized by varying response of the sensor, generating partial overlapping between modalities. Rahman and Csaplovics [36] stressed that such methods have been developed under certain assumptions, and when these assumptions are violated, the result may be of poor quality.

4) Local Dissimilarities Not Taken Into Account: These fusion methods follow global approaches, i.e., the same model applies to the entire image. In the example presented in Fig. 3, with the path visible in the Pan modality and absent in the blue range image, these fusion methods introduce the edges of this path in the fused blue modality. Local dissimilarities between images, such as contrast inversions and occultations of objects, decrease the correlation between the images and are not correctly synthesized by these fusion methods.

In the projection–substitution methods, all images are considered as a vectorial information. Any modification brought by the replacement of $(I_1)_0^{\text{upsamp}}$ by a function of Pan produces a nonquantifiable effect on each modality through the inverse transform. This effect differs according to the adopted projection [7]. It is identical for all MS modalities in the linear IHS fusion method; however, in other color space transforms, a few modalities can concentrate most of the impact.

E. Advantages and Drawbacks of ARSIS Concept Fusion Methods

To limit spectral distortion, the fusion method should preserve low frequencies. It means that if a fused product is downsampled to its original low resolution, the original MS image should be retrieved. This is the consistency property defined by [37]

$$D\left(B_{k1}, (B_{k0}^*)_1\right) < \varepsilon_k \tag{7}$$

where D is the distance between the kth original MS modality B_{k1} and the fused image B_{k0}^* degraded to the original MS resolution *res1*. ε_k is a real value close to zero and depends on the modality. This consistency property captures the first property in [10], [24], and [38]. Multiscale algorithms have the ability to apply hierarchical decomposition of an input image into successive coarser approximations. Such decomposition isolates low frequencies and preserves them while synthesizing high frequencies. Another formulation of this property is possible. If δ is the error between B_l and $(B_h^*)_l$, then

$$(B_{k0}^*)_l = B_{k1} + \delta \tag{8}$$

where δ can be considered as a noise coming from the fusion process and the resampling step. In the Fourier domain, assuming that B_{k1} and δ are uncorrelated, it becomes

$$FT\left((B_{k0}^*)_1\right) = FT(B_{k1}) + FT(\delta) \tag{9}$$

where FT is the Fourier transform. $FT(\delta)$ should be close to zero. Equation (9) means that the frequency content of the original images B_l should be as close as possible to that of fused images degraded to the original resolution l. In other words, fusion methods should only deal with high frequencies.

Many authors already stated that multiscale approaches are able to establish a good tradeoff between the respect of low

frequencies and the insertion of high frequencies [7], [9], [10], [18], [24], [33], [35], [37], [39]. ARSIS concept methods, which are based on multiscale or multiresolution algorithms, inherently fulfill the consistency property.

1) Aliasing: Nevertheless, multiscale algorithms should be chosen in such a way that they do not produce artifacts affecting low frequencies such as aliasing. For instance, the spectral response of the filter used in the HPF method [11] displays bounds yielding to aliasing. Such a filter should not be used.

2) Local Dissimilarities Handling: Several authors developing within the ARSIS concept propose fusion methods based on local approaches or local estimation of parameters to take into account local dissimilarities between images. This is the case of the two couples IMM-HRIMM: CBD (context-based decision) in [20] and RWM (from its authors Ranchin, Wald, Mangolini) in [18]. Basically, the Pan details are injected if the absolute value of the local correlation coefficient is large enough compared to a given threshold, and the MS modalities are upsampled otherwise. The main strength of this approach is to assure good consistency with the original MS data set since nothing is done when images are poorly correlated. However, several experiments demonstrate that it may result into local heterogeneity that degrades the quality of the result and weakens image interpretation.

F. Conclusion

It is worth noticing that another proposal for fusion method classification could have been followed. One of the reviewers proposed a classification into two categories, grouping our second and third categories. He proposed to name them component substitution methods and multiresolution analysis methods, respectively. This classification into two classes permits to distinguish the fusion methods that do not apply a filtering operation from the others. However, the preservation of the spectral content of the original MS data set is ensured only if no disturbing low frequency is injected into the fused products. In our opinion, the preservation of the original content of MS images is very important in the fusion process. It is inherently ensured by the methods in the ARSIS concept. It differentiates this category from the two others, and this is why we favor this classification.

This critical tour of the literature yields better comprehension of the strengths and weaknesses of the different types of methods. Recent methods have been proposed to reduce the artifacts. They are the methods that belong to more than one category and are, thus, called "hybrid methods." Some of them are presented and discussed in Section V.

V. Critical Presentation of Recent Hybrid Methods

A. Projection–Substitution Combined With Relative Spectral Contribution

A group of recent publications has dealt with IHS-based methods combined with the relative spectral contribution assumption. The contribution of each modality to the intensity component is weighted to fit the Pan and, thus, to reduce the radiometric distortion. For instance, Aiazzi *et al.* [40] used the Gram–Schmidt color space transform instead of IHS.

1) IHS and Relative Spectral Contribution Assumption: The first enhancement of the linear IHS transform was made by Tu *et al.* [15]. Particular attention was paid to vegetated areas since spectral distortion is significant in such areas with the original IHS method. The reflectance of vegetation is large in near-infrared range and, thus, in Pan (see, e.g., Fig. 7) and small in visible range. Introducing near-infrared modality into the IHS expression contributes to the increase in the correlation between Pan and this intensity, particularly in presence of vegetation. Similarly to (2), low-resolution intensity is expressed by

$$(I_1)_0^{\text{upsamp}} = 1/4(B_{1,1})_0^{\text{upsamp}} + 1/4(B_{2,1})_0^{\text{upsamp}} + 1/4(B_{3,1})_0^{\text{upsamp}} + 1/4(B_{4,1})_0^{\text{upsamp}} \quad (10)$$

where the fourth modality corresponds to the near-infrared band. The spectral distortion appears to be partially limited. This new method is called the fast IHS and was also found under the name generalized IHS. This equation can easily be generalized to N modalities.

The second enhancement consists of considering the overlapping of the blue and green spectral response functions for the Ikonos imagery (Fig. 7). This overlapping creates spectral redundancy. To decrease this effect, Tu *et al.* [15] modify the previous equation as follows:

$$(I_1)_0^{\text{upsamp}} = 1/3\,[a(B_{1,1})_0^{\text{upsamp}} + b(B_{2,1})_0^{\text{upsamp}} + (B_{3,1})_0^{\text{upsamp}} + (B_{4,1})_0^{\text{upsamp}}] \quad (11)$$

with $a + b = 1$ and $a < b$. These parameters were determined thanks to an analysis on a sample of 92 Ikonos excerpts, maximizing the correlation coefficient $cc()$ between Pan and $(I_1)_0^{\text{upsamp}}$, i.e.,

$$\underset{(a,b)}{\text{Argmax}}\left(cc\left((I_1)_0^{\text{upsamp}}, A\right)\right). \quad (12)$$

An optimal response was obtained for the couple $(a, b) = (0.25, 0.75)$. The end of the procedure remains unchanged. Pan modality replaces $(I_1)_0^{\text{upsamp}}$ before the IHS inverse transform.

Gonzalez-Audicana *et al.* [41] exploited the four-band IHS formulation of Tu *et al.* [15] in (10); however, each MS modality $(B_{k,1})_0^{\text{upsamp}}$ is multiplied by a coefficient γ, which is a function of its spectral response and the Pan one. This coefficient is, thus, different for each modality. All the details of the computation of the weighting factor γ are given in Sections III and IV.

Another version of the new intensity was defined by Choi [42], which is also extended to four MS modalities, i.e.,

$$\text{newI} = A - \left(A - (I_1)_0^{\text{upsamp}}\right)/t \quad (13)$$

where t is the tradeoff parameter to weight Pan injection. For the first time, the tradeoff between spectral and spatial quality

is explicitly introduced. These recent fusion methods agree with Choi [42], who advocated that fusion is a tradeoff modeled by

$$\min_{\text{newI}} \left\{ |A - \text{newI}|^2 + |\text{newI} - (I_1)_0^{\text{upsamp}}|^2 \right\}. \quad (14)$$

2) Discussion: The purpose of these improvements is to reduce the gap between Pan modality and the linear combination of the MS modalities, i.e., to increase the correlation coefficient. Nevertheless, we disagree with this formulation. On the one hand, (14) pushes the new intensity to look like Pan, which is theoretically wrong as discussed in Section II. On the other hand, the second functional term tends to decrease the distance between the new intensity and the original upsampled one. Although the pixel sizes are the same, these two images do not have the same spatial resolution. They are not comparable for high frequencies (see Section II), and it is absolutely wrong to constrain newI to mimic $(I_1)_0^{\text{upsamp}}$.

Equation (14) implies that both high spectral and spatial quality cannot be reached at the same time. Several very recent publications have also dealt with this idea of tradeoff. For instance, Malpica [43] combined generalized IHS with the overlapping treatment of Tu *et al.* [15] with a hue enhancement. Tu *et al.* [44] proposed an improvement of the work of Choi [42] with an automatic adjustment of the t parameter. This idea of tradeoff has also been formulated and used by Lillo-Saavedra and Gonzalo [45], who expressed this tradeoff between the spectral and spatial quality of the fused image using the "a trous" wavelet transform.

We disagree with this idea of tradeoff formulated as an *a priori* assumption for fusion method developments. We think that an ideal fusion method must be able to simultaneously reach both radiometric and geometric quality, and not one at the expense of the other. It is in no way possible to define weights to write Pan as a function of MS modalities with the recent high-resolution satellite sensors. Such assumptions restrict the quality of the results.

The ideal fusion method should be able to preserve original spectral and spatial information of the MS images while increasing the spatial resolution. Dissociating the spectral and spatial quality has drastic consequences. Many authors think that the color information belongs to the original MS images, whereas the high spatial resolution details belong to the Pan modality. This idea led [45] to the design of a fusion method that is based on the minimization of a distance between the Pan modality and each MS band, as if high-frequency details should be exactly those of the Pan, which is often locally false as illustrated in previous sections. The important point here is that high-resolution details also bear spectral information. A truck appearing in bright red at low resolution should remain red in the high-resolution fused color composition image, which means that high-frequency details should be reduced to the red modality in this particular case. That is why spectral and spatial aspects of fusion are intrinsically linked. As the ideal fusion method does not exist yet, it turns out that the fused products, according to the tool used for fusion, generally correspond to a tradeoff between a good geometrical representation of structures and a good representation of original colors. Therefore, according to our opinion and experience, this tradeoff is a consequence of the use of certain tools and not an *a priori* compulsory assumption for development.

B. Relative Spectral Contribution Combined With the ARSIS Concept Assumption

Several recent fusion methods are based on the minimization of energy functionals E. These are hybrid methods because the terms of the functionals are selected in different categories of methods. For instance, Ballester *et al.* [46] proposed an algorithm based on the following terms and assumptions.

- The geometry of the MS images is contained in the topographic map of the Pan modality, which defines the term $E_{\text{geometric}}$. This term uses a gradient expression to extract high frequencies from the Pan and to inject them into each MS modality, and is, thus, in accordance with the ARSIS concept.
- There exists a linear relationship between the Pan and the MS images, giving $E_{\text{radiometric}}$, like in the case of relative spectral contribution methods.
- The last term corresponds to the expression of the low-resolution pixel function of the high-resolution pixels using a smoothing kernel $E_{\text{linkedTOdata}}$. This is the mathematical translation of the consistency property introduced in the presentation of the ARSIS concept.

These three elements build an energy functional whose minima give the reconstructed MS images at higher resolution. This functional is written as follows:

$$E_k = \mu E_{\text{linkedTOdata}} + E_{\text{geometric}} + \lambda E_{\text{radiometric}} \quad (15)$$

where μ and λ are positive input parameters weighting the proportion of each term. E_k is the energy for the kth modality to be fused. The minimization can be performed using the following gradient descent:

$$((B_{k1})_0^*)^{p+1} = ((B_{k1})_0^*)^p + t(\nabla E_k)^p \quad (16)$$

with t the parameter driving the gradient descent and p the index of iteration; therefore, $(B_{k0}^*)^p$ is the kth fused modality at the pth iteration, and $(B_{k0}^*)^{p+1}$ is this same modality at the next iteration. The final image results from a tradeoff between the constraints modeled by the various terms. This tradeoff depends on the real input parameters (μ and λ). The radiometric term constrains the fused modalities to fit the linear combination between Pan and MS at the high spatial resolution in the iterative process, whereas, as previously discussed, it is not exactly true even at the low resolution. We experimented this method and found it very performing in urban areas, where the correlation between MS modalities and Pan is high. However, in case of natural landscapes, the algorithm fails and introduces a local distortion in the colors. Even with the weighting factors for the ARSIS terms, the radiometric assumption reduces the fusion performance.

We also paid attention to another functional with a new *a priori* assumption, which is developed in [47]. Energy also articulates around the sum of three terms, shown as follows:

$$E_k = \alpha_1 E_{\text{linkedTOdata}} + \alpha_2 E_{\text{physics}} + \alpha_3 E_{\text{smoothness}} \quad (17)$$

where α_1, α_2, and α_3 are positive weighting parameters. Minimization is performed with the Metropolis algorithm. $E_{\mathrm{linkedTOdata}}$ is similar to the one proposed in [46]. It enforces the spectral consistency between the original low-resolution MS image and the synthesized high-resolution image. This is achieved using a smoothing filter to degrade the spatial resolution. Since a multiscale assumption is driving this term, it can be related to an ARSIS concept method. Similarly, E_{physics} reminds the geometric term in [46]; however, the variance is used instead of the gradient.

The term of smoothness is the most problematic. It was originally defined to struggle against the blocky effect induced by the technical constraints in implementing $E_{\mathrm{linkedTOdata}}$. However, such a term dramatically limits innovation. Small-scale variations and features that should appear at the highest spatial resolution are smoothed away by this term, which decreases visual quality. Authors were aware of it and wrote, "this could simply imply that the smoothness constraint [...] was not a good *a priori* assumption about the ground truth."

C. Projection–Substitution Combined With the ARSIS Concept Assumption

We recall that the spectral distortion in projection–substitution methods is caused by a modification of low frequencies of each original MS. Several authors combine IHS and multiscale approaches such as wavelet transforms. References [48]–[51] used the wavelet transform to select only high frequencies of Pan to inject into $(I_1)_0^{\mathrm{upsamp}}$. The advantage of this approach is to respect the consistency property [10], [24], [37], as it preserves low frequencies of the original MS images by preserving the low frequencies of $(I_1)_0^{\mathrm{upsamp}}$. The colors of fused products are faithful to the original color composite images. Moreover, the visual impression of the synthesized image is very good since the whole structures of Pan are introduced.

However, the existence of dissimilarities, as illustrated in Section II, is not tackled by any of these models. None of these methods uses a local approach to check whether a structure should appear in the new intensity or fused modalities. If one refers to the example of Fig. 2, the path will appear in the fused images.

In conclusion, recent hybrid developments, which exploit both projection–substitution and relative spectral contribution principles, tend to gather these two categories into a unique one. In both cases, the fused product $(B_{k1})_0^*$ is a function of all original Pan and MS images. The only difference is that projection–substitution methods synthesize all MS modalities at the same time and not separately as for the relative spectral contribution ones.

VI. Conclusion

A fused product reaches a good quality only if the characteristics and differences between input images are taken into account. Figures in this paper show that dissimilarities existing between these two data sets originate from two causes—different times and different spectral bands of acquisition. Remote sensing physics should be carefully considered while designing the fusion process. Because of the complexity of physics and the large number of unknowns, authors are led to make assumptions to drive their development.

Weaknesses and strengths of each reported method have been discussed. The conclusion of this critical survey of literature is that the choice in the assumptions for the development of a method should be carefully done, with the risk to drastically weaken fusion performance in certain situations.

The higher the correlation between Pan and each MS modality, the better the success of projection–substitution and relative spectral contribution methods. By construction, these methods introduce low frequencies into an MS modality coming from other modalities. The consequence is a radiometric distortion whose importance is linked to this correlation coefficient value.

Even with the last efforts made to decrease the spectral distortion of these two types of methods by introducing the ARSIS concept assumption, this artifact cannot completely disappear if it keeps on exploiting the linear combination linking the Pan modality to all MS ones, even to a smaller extent. The tradeoff between spectral and spatial quality is not a fatality, and we think that an ideal fusion method should be able to simultaneously reach quality in both domains.

This critical survey of literature leads us to promote approaches using multiscale or multiresolution algorithms, i.e., fusion developments made within the ARSIS concept. Such algorithms split the high from the low frequencies. Provided the choice of an efficient multiscale algorithm avoiding aliasing, the consistency property is checked, and the spectral distortion is limited.

The remaining artifacts are linked to a bad synthesis of the MS high frequencies. A global synthesis does not appropriately take local occultations and contrast inversions into account. Such defaults of global ARSIS methods are also limited if the correlation coefficient between Pan and each MS modality is high. Several works demonstrate that a local approach is capable of handling these effects. However, it may generate local heterogeneity in the fused images. A possible solution lies in the combination of local and global approaches [52], which will produce fused images with a tradeoff between the advantages and drawbacks of the two approaches.

Among the local dissimilarities raised in Section II, the most unknown is certainly that illustrated in Fig. 3, when an object is perceivable in an MS modality, whereas it is not visible in the Pan image. In this particular situation, high-resolution MS details of this object cannot be inferred from the Pan modality. The only available information concerning this object belongs to the considered MS image. Therefore, the ideal fusion method should be able to refer to the intrapyramidal MS information, coming from lower scales, to extrapolate the high details of this particular structure.

In summary, we think that the methods calling upon the ARSIS concept are the only way to tend toward the ideal fusion method. Results obtained up to now by these methods are not perfect; nevertheless, they are at least as good as those from

other methods [53]. The major advantage compared to the other types of methods is that the ARSIS concept is a framework respecting the physical properties of the original MS images that are prevented from large spectral distortion and still offers a large number of opportunities for development.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers whose comments helped in improving the clarity and quality of this paper. The benefit of discussions held in the framework of the working group "data fusion" of the European Association of Remote Sensing Laboratories is acknowledged.

REFERENCES

[1] L. Wald, "Some terms of reference in data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1190–1193, May 1999.

[2] G. Piella, "A general framework for multiresolution image fusion: From pixels to regions," *Inf. Fusion*, vol. 4, no. 4, pp. 259–280, Dec. 2003.

[3] A. Filippidis, L. C. Jain, and N. Martin, "Multisensor data fusion for surface land-mine detection," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 30, no. 1, pp. 145–150, Feb. 2000.

[4] P. S. Huang and T. Te-Ming, "A target fusion-based approach for classifying high spatial resolution imagery," in *Proc. Workshops Advances Techn. for Analysis Remotely Sensed Data*, Oct. 27–28, 2003, pp. 175–181.

[5] J. Inglada, V. Muron, D. Pichard, and T. Feuvrier, "Analysis of artifacts in subpixel remote sensing image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 254–264, Jan. 2007.

[6] P. Blanc, L. Wald, and T. Ranchin, "Importance and effect of co-registration quality in an example of 'pixel to pixel' fusion process," in *Proc. 2nd Conf. Fusion Earth Data: Merging Point Meas., Raster Maps Remotely Sensed Images*, T. Ranchin and L. Wald, Eds., Nice, France: SEE/URISCA, Jan. 28–30, 1998, pp. 67–73.

[7] J. Zhou, D. L. Civco, and J. A. Silvander, "A wavelet transform method to merge Landsat TM and SPOT panchromatic data," *Int. J. Remote Sens.*, vol. 19, no. 4, pp. 743–757, Mar. 1998.

[8] Y. Zhang, "A new merging method and its spectral and spatial effects," *Int. J. Remote Sens.*, vol. 20, no. 10, pp. 2003–2014, Jul. 1999.

[9] T. Ranchin and L. Wald, "Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation," *Photogramm. Eng. Remote Sens.*, vol. 66, no. 1, pp. 4–18, 2000.

[10] L. Wald, *Data Fusion: Definitions and Architectures. Fusion of Images of Different Spatial Resolutions*. Paris, France: Les Presses de l'école des Mines, 2002, p. 197.

[11] P. S. Chavez, S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 3, pp. 265–303, 1991.

[12] R. Haydn, G. W. Dalke, J. Henkel, and J. E. Bare, "Application of the IHS color transform to the processing of multisensor data and image enhancement," in *Proc. Int. Symp. Remote Sens. Arid, Semi-Arid Lands*, Cairo, Egypt, 1982, pp. 599–616.

[13] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Reading, MA: Addison-Wesley, 1992.

[14] T. M. Tu, S. C. Su, H. C. Shyu, and P. S. Huang, "A new look at IHS-like image fusion methods," *Inf. Fusion*, vol. 2, no. 3, pp. 177–186, Sep. 2001.

[15] T. M. Tu, P. S. Huang, C. L. Hung, and C. P. Chang, "A fast intensity-hue-saturation fusion technique with spectral adjustment for Ikonos imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 309–312, Oct. 2004.

[16] A. R. Gillespie, A. B. Kahle, and R. E. Walker, "Color enhancement of highly correlated images—II Channel ratio and 'chromacity' transformation techniques," *Remote Sens. Environ.*, vol. 22, no. 3, pp. 343–365, Aug. 1987.

[17] Guide des Utilisateurs des Données SPOT. Anonymous, 1986., CNES and SPOT Image (Eds.), Toulouse, France, 3 volumes, revised January 1991.

[18] T. Ranchin, B. Aiazzi, L. Alparone, S. Baronti, and L. Wald, "Image fusion—The ARSIS concept and some successful implementation schemes," *ISPRS J. Photogramm. Remote Sens.*, vol. 58, no. 1/2, pp. 4–18, Jun. 2003.

[19] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2300–2312, Oct. 2002.

[20] D. Pradines, "Improving SPOT image size and multispectral resolution," in *Proc. SPIE Conf. Earth Remote Sens. Using Landsat Thematic Mapper SPOT Syst.*, 1986, vol. 660, pp. 78–102.

[21] J. G. Liu and J. McMoore, "Pixel block intensity modulation: Adding spatial detail to TM band 6 thermal imagery," *Int. J. Remote Sens.*, vol. 19, no. 13, pp. 2477–2491, Sep. 1998.

[22] J. C. Price, "Combining multispectral data of differing spatial resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1199–1203, May 1999.

[23] J. G. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, Dec. 2000.

[24] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogramm. Eng. Remote Sens.*, vol. 63, no. 6, pp. 691–699, Jun. 1997.

[25] P. Blanc, T. Blu, T. Ranchin, L. Wald, and R. Aloisi, "Using iterated rational filter banks within ARSIS concept for producing 10 m Landsat multispectral images," *Int. J. Remote Sens.*, vol. 19, no. 12, pp. 2331–2343, Aug. 1998.

[26] S. Ioannidou and V. Karathanassi, "Investigation of the dual-tree complex and shift-invariant discrete wavelet transforms on Quickbird Image Fusion," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 1, pp. 166–170, Jan. 2007.

[27] M. Choi, R. Y. Kim, M.-R. Nam, and H. O. Kim, "Fusion of multispectral and panchromatic satellite images using the curvelet transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 2, pp. 136–140, Apr. 2005.

[28] H. Song, S. Yu, L. Song, and X. Yang, "Fusion of multispectral and panchromatic satellite images based on contourlet transform and local average gradient," *Opt. Eng.*, vol. 46, 020 502, 2007.

[29] C. Thomas and L. Wald, "A MTF-based distance for the assessment of geometrical quality of fused products," in *Proc. Fusion (IEEE Catalog 06EX1311C)*, Florence, Italy, Jul. 10–13, 2006, pp. 1–7. (CD-ROM).

[30] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "An MTF-based spectral distortion minimizing model for pan-sharpening of very high resolution multispectral images of urban areas," in *Proc. 2nd GRSS/ISPRS Joint Workshop Remote Sens. Data Fusion Over Urban Areas*, Berlin, Allemagne, May 22–23, 2003, pp. 90–94. IEEE Ed. (03EX646).

[31] V. K. Shettigara, "A generalized component substitution technique for spatial enhancement of multispectral images using higher resolution data set," *Photogramm. Eng. Remote Sens.*, vol. 58, no. 5, pp. 561–567, 1992.

[32] V. Vijayaraj, C. O'Hara, and N. Younan, "Quality analysis of pan-sharpened images," in *Proc. IEEE IGARSS*, 2004, vol. 1, pp. 85–88.

[33] D. A. Yocky, "Multiresolution wavelet decomposition image merger of Landsat Thematic Mapper and SPOT panchromatic data," *Photogramm. Eng. Remote Sens.*, vol. 62, no. 9, pp. 1067–1074, 1996.

[34] A. Pellemans, R. Jordans, and R. Allewiijn, "Merging multispectral and panchromatic SPOT images with respect to the radiometric properties of the sensor," *Photogramm. Eng. Remote Sens.*, vol. 59, no. 1, pp. 81–87, 1993.

[35] W. Shi, C. Zhu, Y. Tian, and J. Nichol, "Wavelet-based image fusion and quality assessment," *Int. J. Appl. Earth Observation Geoinformation*, vol. 6, no. 3/4, pp. 241–251, Mar. 2005.

[36] M. M. Rahman and E. Csaplovics, "Examination of image fusion using synthetic variable ratio (SVR) technique," *Int. J. Remote Sens.*, vol. 28, no. 15, pp. 3413–3424, Jan. 2007.

[37] C. Thomas and L. Wald, "Assessment of the quality of fused products," in *Proc. 24th EARSeL Symp. New Strategies for Eur. Remote Sens.*, Dubrovnik, Croatia, Oluic, Ed. Rotterdam, The Netherlands: Millpress, May 25–27, 2004, pp. 317–325.

[38] J. Li, "Spatial quality evaluation of fusion of different resolution images," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 33, no. B2-2, pp. 339–346, 2000.

[39] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, "A comparative analysis of image fusion methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 81–84, Jun. 2005.

[40] B. Aiazzi, L. Alparone, S. Baronti, and M. Selva, "MS + Pan image fusion by enhanced Gram-Schmidt spectral sharpening," in *Proc. 26th EARSeL Symp. New Strategies for Eur. Remote Sens.*, Warsaw, Poland, May 29–31, 2006.

[41] M. Gonzalez-Audicana, X. Otazu, O. Fors, and J. Alvarez-Mozos, "A low computational-cost method to fuse Ikonos images using the spectral response function of its sensors," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1683–1691, Jun. 2006.
[42] M. Choi, "A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1672–1682, Jun. 2006.
[43] J. A. Malpica, "Hue adjustment to IHS Pan-sharpened Ikonos imagery for vegetation enhancement," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 1, pp. 27–31, Jan. 2007.
[44] T.-M. Tu, W.-C. Cheng, C.-P. Chang, P. S. Huang, and J.-C. Chang, "Best tradeoff for high-resolution image fusion to preserve spatial details and minimize color distortion," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 2, pp. 302–306, Apr. 2007.
[45] M. Lillo-Saavedra and C. Gonzalo, "Spectral or spatial quality for fused satellite imagery? A trade-off solution using the wavelet a trous algorithm," *Int. J. Remote Sens.*, vol. 27, no. 7, pp. 1453–1464, Apr. 2006.
[46] C. Ballester, V. Caselles, J. Verdera, and B. Rougé, "A variational model for P + XS image fusion," in *Proc. IEEE Workshop Variational, Geometric Level Sets Method Comput. Vision*, Oct. 12, 2003.
[47] A. Vesteinsson, A. Henrik, J. R. Sveinsson, and J. A. Benediktsson, "Spectral consistent satellite image fusion: Using a high resolution panchromatic and low resolution multi-spectral images," in *Proc. IEEE IGARSS*, Seoul, Korea, 2005, vol. 4, pp. 2834–2837.
[48] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Palà, and R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1204–1211, May 1999.
[49] Y. Chibani and A. Houacine, "The joint use of IHS transform and redundant wavelet decomposition for fusing multispectral and panchromatic images," *Int. J. Remote Sens.*, vol. 23, no. 18, pp. 3821–3833, Sep. 2002.
[50] M. Gonzalez-Audicana, J. L. Saleta, O. G. Catalan, and R. Garcia, "Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1291–1299, Jun. 2004.
[51] Z. Li, Z. Jing, X. Yang, and S. Sun, "Color transfer based remote sensing image fusion using non-separable wavelet frame transform," *Pattern Recognit. Lett.*, vol. 26, no. 13, pp. 2006–2014, Oct. 2005.
[52] C. Thomas, "Fusion d'images de résolutions spatiales différentes," Thèse de Doctorat en Mathématiques appliquées et traitement d'images, Ecole des mines de Paris, Paris, France, 2006.
[53] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.



Claire Thomas received the B.S. degree in physics from the Ecole Nationale Supérieure de Physique de Strasbourg, Illkirch, France, and the M.S. degree in photonics, image, and cybernetics from the University of Strasbourg, Strasbourg, France, both in 2002, and the Ph.D. degree from the Ecole des Mines de Paris, Sophia Antipolis, France, in 2006.

She is currently with the Ecole des Mines de Paris/Armines. By far, she has presented seven communications and published two peer-reviewed articles.



Thierry Ranchin (M'01) received the Ph.D. degree in the field of applied mathematics from the University of Nice Sophia Antipolis, Nice, France, in 1993 and the "Habilitation à diriger les recherches," in 2005.

After being a Postdoctoral Fellow in a company in Tromso, Norway, he joined the remote sensing group of the Ecole des Mines de Paris/Armines, Sophia Antipolis, France, in the fall of 1994. He was an invited scientist from the University of Jena, Jena, Germany, in 1998. He is the Cochair of the Energy Community of Practices of the Global Earth Observation System of Systems initiative. Since January 2007, he has been the Head of the Observation, Modeling, and Decision Team of the Center for Energy and Processes, Ecole des Mines de Paris. He is the holder of a patent about sensor fusion and has published more than 100 publications, communications in international symposia, chapters of books, or articles in journals with peer-review committees in the field of remote sensing of the Earth system and in the field of image processing. He was the Coeditor of the series of conferences "Fusion of Earth Data."

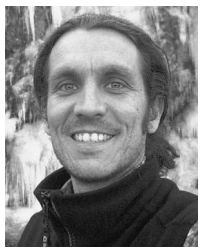
Dr. Ranchin is the recipient of the Autometrics Award in 1998 and the Erdas Award in 2001 from the American Society for Photogrammetry and Remote Sensing for articles on data fusion.



Lucien Wald received the B.S. degree in theoretical physics in Marseille and Paris, France, in 1977, the Ph.D. degree on the applications of remote sensing to oceanography in Paris in 1980, and the Doctorat d'Etat ès Sciences on the applications of remote sensing to oceanography in Toulon, France, in 1985.

Since 1991, he has been a Professor with the Ecole des Mines de Paris, Sophia Antipolis, France. He is focusing his own research in applied mathematics and meteorology.

Dr. Wald is the recipient of the Autometrics Award in 1998 and the Erdas Award in 2001 for articles on data fusion. His career in information technologies has been rewarded in 1996 by the famous French Blondel Medal.



Jocelyn Chanussot (M'04–SM'04) received the B.S. degree in electrical engineering from the Grenoble Institute of Technology (INP Grenoble), Grenoble, France, in 1995 and the Ph.D. degree from the University of Savoie, Annecy, France, in 1998.

In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement. Since 1999, he has been teaching signal and image processing with the INP Grenoble and has been working at the Grenoble Images Speech Signals and Automatics Laboratory, Grenoble, as an Assistant in 1999–2005, an Associate in 2005–2007, and, currently, a Full Professor. His research interests include statistical modeling, multicomponent image processing, nonlinear filtering, remote sensing, and data fusion.

Dr. Chanussot is an Associate Editor for the journal *Pattern Recognition* (2006–2008). He was an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (2005–2007) and is currently an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (2007–2010). He was the Cochair of the IEEE Geoscience and Remote Sensing Society Data Fusion Technical Committee (2005–2007) and a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006–2008).