



## A Variational Model for P+XS Image Fusion

COLOMA BALLESTER, VICENT CASELLES, LAURA IGUAL AND JOAN VERDERA

*University Pompeu Fabra, Departament de Tecnologia, Pg. Circumvalació 8, 08003 Barcelona, Spain*

coloma.ballester@upf.edu

vicent.caselles@upf.edu

laura.igual@upf.edu

joan.verdera@upf.edu

BERNARD ROUGÉ

*CNES, Av. Ed Belin, 31401 Toulouse, France*

bernard.rouge@cnes.fr

*Received March 11, 2004; Revised May 26, 2005; Accepted October 26, 2005*

*First online version published in April, 2006*

**Abstract.** We propose an algorithm to increase the resolution of multispectral satellite images knowing the panchromatic image at high resolution and the spectral channels at lower resolution. Our algorithm is based on the assumption that, to a large extent, the geometry of the spectral channels is contained in the topographic map of its panchromatic image. This assumption, together with the relation of the panchromatic image to the spectral channels, and the expression of the low-resolution pixel in terms of the high-resolution pixels given by some convolution kernel followed by subsampling, constitute the elements for constructing an energy functional (with several variants) whose minima will give the reconstructed spectral images at higher resolution. We discuss the validity of the above approach and describe our numerical procedure. Finally, some experiments on a set of multispectral satellite images are displayed.

**Keywords:** multispectral images, topographic map, image fusion, energy functional

### 1. Introduction

A grey-level image can be realistically modeled as a real function  $u(x)$  where  $x$  represents an arbitrary point of a rectangle  $\Omega$  in  $\mathbb{R}^2$  and  $u(x)$  denotes the grey level at  $x$ . Typically  $u(x)$  represents the photonic flux over a wide band of wavelengths and we have a proper grey-level image. Below, we shall refer to it as the panchromatic image. A multispectral image may be represented by a function  $\vec{u}$  from  $\Omega$  to  $\mathbb{R}^m$  where  $m$  represents the number of spectral channels. For colour images, typically,  $m = 3$  if we consider the usual  $R$ ,  $G$ ,  $B$  channels. If we add the near infrared channel, we have a

multispectral image with  $m = 4$ . In this case, each coordinate of  $\vec{u}(x)$  represents the intensity corresponding to a spectral channel, where the photonic flux is subjected to a spectral selective filter, be it in the visible range, the near infrared or the ultraviolet. Most satellites provide us with multispectral images which include the usual color channels plus the infrared one, and/or the ultraviolet one (see <http://smc.cnes.fr>). Let us mention here that the abbreviation  $P + XS$  stands for Panchromatic and MultiSpectral images.

We say that the panchromatic image  $u$  corresponds to the multispectral image  $\vec{u}$  if  $u(x)$  has been obtained by adding (with some mixing coefficients) the coordinates

of  $\vec{u}$ , which represent the energies of different spectral bands. In other words, if  $\vec{u} = (u_1, \dots, u_m)$  and  $\alpha_1, \dots, \alpha_m$  represent the mixing coefficients for computing the total spectral energy from the above channels, the corresponding panchromatic image will be given by

$$u(x) = \alpha_1 u_1(x) + \dots + \alpha_m u_m(x). \quad (1)$$

In digital images, the only accessible information is a sampled and quantized version of  $u$ ,  $u(i, j)$ , where  $(i, j)$  is a set of discrete points (in general, on a grid) and  $u(i, j)$  belongs to a discrete set of values,  $0, 1, \dots, 255$ , in many cases. Since, by Shannon's theory, we can assume that  $u(x)$  is recoverable at any point from the samples  $u(i, j)$ , as a first approximation, we may assume that the image  $u(x)$  is known in a continuous domain, up to the quantization noise.

This paper describes a method of increasing the resolution of satellite multispectral images when we know the corresponding panchromatic image  $u$  at a higher resolution. We assume that the panchromatic image  $u(i, j)$  has a size of  $N \times N$  pixels. The spectral channels have been sampled at a lower resolution giving images of size  $\frac{N}{s} \times \frac{N}{s}$  (typically  $s = 2$ , or  $4$ ), which we shall denote by  $\vec{u}^s = (u_1^s, \dots, u_m^s)$ . The superscript  $s$  is explicitly included to stress the loss of resolution of the multispectral data. In satellite imaging, acquiring the panchromatic image at high resolution and subsampling the multispectral channels by a factor of 2, or 4, permits a substantial reduction of the data to be transmitted to the earth, but it leaves us with the problem of recovering them at high resolution. Our purpose will be to reconstruct the high-resolution multispectral image  $\vec{u} = (u_1, \dots, u_m)$ , knowing the data  $u$  and  $\vec{u}^s$ , and taking into account several constraints imposed by the data generation model. The first constraint to be satisfied is given by relation (1). The other ones express the fact that, for each channel, the low-resolution pixels are formed from the high-resolution ones by low-pass filtering followed by subsampling. If we denote by  $k_r$  the impulse response of the filter corresponding to channel  $r \in \{1, \dots, m\}$ , then we may write these constraints as

$$u_r^s(i, j) = k_r * u_r(si, sj) \quad \forall i, j \in \left\{0, \dots, \frac{N}{s} - 1\right\}. \quad (2)$$

The values of  $\alpha_i$  and the kernels  $k_r$ ,  $i, r = 1, \dots, m$ , depend on the type of satellite (see <http://smc.cnes.fr>).

Finally, we use the geometric information contained in the panchromatic image. Indeed, we assume in this paper that the panchromatic image and the low-resolution multispectral images are aligned and we constrain the geometry of the spectral channels at high resolution to follow the geometry of the panchromatic image. This constitutes the main feature of our approach and needs further explanation. First, in Section 1.1 we explain what we mean by the *geometric information* contained in the panchromatic image, and more generally, in any scalar image, including any spectral channel. Then, in Sections 1.2 and 1.3, we explain the reasons supporting the underlying assumption that *images taken on different spectral bands share common geometric information*. For that, we shall review the main conclusions in Caselles et al. (2002), and Lisani and Morel (2000). The rest of the paper is devoted to present our model and display some numerical experiments.

The literature on the subject is extensive. Most methods are based on the injection of high frequency components (corresponding to spatial details present in the high resolution panchromatic image) in interpolated versions of the multispectral data (Wald et al., 1997), or improvements based on multiresolution analysis of P+XS images (Aiazzi et al., 2002). For more information we refer to Aiazzi et al. (2002) and Wald et al. (1997), and references therein.

### 1.1. *Mathematical Morphology of Scalar Images*

This subsection considers scalar images, that is, images with a single channel, either a color (or any other spectral channel) or grey level.

We now consider the question: what is the geometric information content of an image? This section summarizes some arguments contained explicitly or implicitly in Mathematical Morphology theory (Serra, 1982), which were further developed in Caselles et al. (1999).

The sensors of a camera or a CCD array transform the continuum of light energies to a finite interval of values by means of a nonlinear contrast function  $g$ . The contrast change  $g$  depends on the properties of the sensors, but also on the illumination conditions and the reflection properties of the objects, and those conditions are generally unknown. Images are observed modulo an arbitrary and unknown contrast change. These observations lead the physicist and psychologist M. Wertheimer (Wertheimer, 1923) to state, as a basic principle, that the gray level is not an observable.

Mathematical Morphology recognized contrast invariance as a basic requirement and proposed that image analysis operations take into account this invariance principle (Serra, 1982; Guichard and Morel, 2003). With this principle, an image  $u$  is a representative of an equivalence class of images  $v$  obtained from  $u$  via a contrast change, i.e.,  $v = g(u)$  where  $g$ , for simplicity, will be a continuous strictly increasing function. Under this assumption, an image is characterized by its upper (or lower) level sets  $X_\lambda = [u \geq \lambda] = \{x : u(x) \geq \lambda\}$  (resp.  $X'_\lambda = [u \leq \lambda] = \{x : u(x) \leq \lambda\}$ ). Moreover, the image can be recovered from its level sets by the reconstruction formula

$$u(x) = \sup\{\lambda : x \in X_\lambda\}.$$

Thus, according to the Mathematical Morphology doctrine, the reliable information of the image is contained in the level sets, independent of their actual levels. Thus, we are led to consider that the geometric information, the shape information, is contained in the family of level sets of the image.

We can further describe the level sets by their boundaries,  $\partial X_\lambda u$ , which are, under suitable, very general assumptions, Jordan curves (Jordan curves are continuous maps from the circle into the plane  $\mathbb{R}^2$  without crossing points). Indeed, in Ambrosio et al. (2001) it is proved that if  $u$  is a function whose upper level sets  $X_\lambda u$  are sets of finite perimeter (in particular, if  $u$  is a function of bounded variation (Ambrosio et al., 2000; Evans and Gariepy, 1992)), then the boundaries of level sets can be described by a countable family of Jordan curves with finite length. The family of all level lines of an image was called the *topographic map* (Caselles et al., 1999). The topographic map is invariant under a wide class of local contrast changes (Caselles et al., 1999), and, in particular, it is a useful tool for comparing images of the same object with different illuminations (Lisani and Morel, 2000; Monasse, 1999). We can conceive the topographic map as a tool giving a complete description of the geometry of grey-level images.

For later purposes, we have to go a step further in the description of the geometry of the topographic map. Assuming that our image  $u : \Omega \rightarrow \mathbb{R}$  is of bounded variation, by the results in Ambrosio et al. (2001), almost all its level sets  $[u \geq \lambda]$  are sets of finite perimeter whose boundaries are unions of rectifiable Jordan curves, and we may compute the unit normal vector

at almost any point of it. Thus, if  $x$  is on the essential boundary of an upper level set  $[u \geq \lambda]$ ,  $\lambda \in \mathbb{R}$ , we define  $\theta(x)$  as the unit normal to the essential boundary of  $[u \geq \lambda]$ . On the set of points where the vector field  $\theta$  is defined, the identity  $\theta \cdot \nabla u = |\nabla u|$  holds (in a rigorous mathematical sense this identity has to be understood as the equality of two measures). The vector field  $\theta$  has to be extended to all of  $\Omega$ . For that, we define  $\theta(x) = \frac{\nabla u(x)}{|\nabla u(x)|}$  if  $\nabla u(x) \neq 0$  and  $\theta(x) = 0$  when  $\nabla u(x) = 0$ . In other words, in a flat region of the graph of  $u$  we define  $\theta = 0$ , and at the points  $x \in \Omega$  which are on the essential boundaries of the upper level sets of  $u$  we define  $\theta(x)$  as the unit normal vector field. A more precise description of  $\theta$  will be given in Section 2 where the word “essential”, when referring to boundaries, will be explained. This vector field will be the analytic tool we use to impose the constraint that the geometry of a function  $v$  can be described in terms of the geometry of  $u$ , i.e., the level lines of  $v$  are a subset of the family of level lines of  $u$ .

## 1.2. Geometry and Color in Natural Images

What is the geometric content of a color image? Obviously, the answer to this question is quite complex and, strictly speaking, it cannot be reduced to the geometry of its associated intensity image. Indeed, counterexamples can be given where color objects exist with a constant intensity. But what happens in images of natural scenes? Will the light create color patterns with color edges with a constant intensity? (Caselles et al., 2002) experimentally confirmed that the essential geometric contents of a color image is contained in the level lines of the corresponding intensity image.

To support the above conclusion, the authors of Caselles et al. (2002) proposed an algorithm to constrain the color channels of a given image to have the same geometry (i.e. the same level lines) as the grey level. The algorithm *replaces the colors in an image by their conditional expectation with respect to the grey levels*. If the hypothesis above is sound, then this algorithm should not alter the colors of the image or its visual aspect. The authors of Caselles et al. (2002) displayed several experiments confirming this hypothesis. A further confirmation of it was suggested by the experiments of imposing the color of one image on the topographic map of another one: It resulted, in a striking way, in the dominance of grey level and the fading of a color deprived of its geometry (Caselles et al., 2002).

### 1.3. The Case of Satellite Multispectral Images

In the case of multispectral satellite images, the analogous assumption would be that images of the same scene taken on different wavelength bands would share common geometric information. Indeed, this assumption was experimentally studied in Lisani and Morel (2000). The experiments were done on images of four spectral channels corresponding to the blue, green, red and near infrared regions. Two channels were compared via their topographic maps. Two types of comparison were proposed: by means of the unit normal vector field of the topographic map, and by means of pieces of its level lines (Lisani and Morel, 2000). With both comparison procedures, the conclusion was the same: the channels that are in the visible region share a large portion of its topographic map. The amount of sharing decreases, but is still large, for the near infrared channel. For the red and near infrared images, it was also shown that, after contrast inversion, there is still a portion of the topographic map that is common to the topographic map of the blue channel (Lisani and Morel, 2000). This was explained in terms of the vegetation, which becomes apparent in the red and near infrared channels by contrast inversion (Lisani and Morel, 2000). The methods used were robust under illumination changes and geometric transformations, in particular, under changes of scale (Lisani and Morel, 2000).

### 1.4. Conclusion

Thus, based on the above arguments, we shall adopt the hypothesis that for satellite multispectral images, *to a large extent, the geometry of the spectral channels is contained in the topographic map of its panchromatic image.*

This assumption, together with (1) and (2), is the basis of our variational approach to increasing the resolution of satellite multispectral images when we know the corresponding panchromatic image at a higher resolution.

Section 2 recalls some basic facts about functions of bounded variation. Section 3 describes the energy functional of our variational model. Section 4 covers the algorithm and the numerical experiments. Section 5 presents some conclusions. Finally, Section 5 briefly comments on the mathematical justification of the problem.

## 2. Function Spaces

This section contains some mathematical definitions that are needed for proper definitions of the energy functionals below. It may be skipped, since the main notations will be recalled in a simpler way in Section 3.

Let  $Q$  be an open set in  $\mathbb{R}^2$ . We denote by  $L^p(Q)$  the space of functions  $u : Q \rightarrow \mathbb{R}$  such that  $|u|^p$  is Lebesgue integrable. We denote by  $C_0^\infty(Q)$  the space of infinitely differentiable functions in  $Q$  with compact support (contained in  $Q$ ). By  $C(\bar{Q})$  we denote the space of continuous functions in  $\bar{Q}$  (which denotes the closure of  $Q$ ).

Let us recall the definition of  $BV$  functions and total variation. A function  $u \in L^1(Q)$  whose partial derivatives, in the sense of distributions, are measures with finite total variation in  $Q$  is called a function of bounded variation. The class of such functions will be denoted by  $BV(Q)$ . Thus  $u \in BV(Q)$  if and only if there are Radon measures  $\mu_1, \mu_2$  defined in  $Q$  with finite total mass in  $Q$  and

$$\int_Q u D_i \varphi dx = - \int_Q \varphi d\mu_i \quad (3)$$

for all  $\varphi \in C_0^\infty(Q)$ ,  $i = 1, 2$ . Thus, the gradient of  $u$ , denoted by  $\nabla u$ , is a vector-valued measure with finite total variation

$$\|\nabla u\| = \sup \left\{ \int_Q u \operatorname{div} \varphi dx : \varphi \in C_0^\infty(Q, \mathbb{R}^2), \right. \\ \left. |\varphi(x)| \leq 1 \text{ for } x \in Q \right\}. \quad (4)$$

If  $\nabla u$  is an integrable function, then

$$\|\nabla u\| = \int_Q |\nabla u| dx. \quad (5)$$

In any case, if  $u \in BV(Q)$  we shall always write  $\|\nabla u\| = \int_Q |\nabla u| dx$ . At this point, we notice that we are not using the standard notation that does not incorporate the  $dx$ , since  $\nabla u$  is a measure. This choice permits to unify the notation in the formulas of Section 3. The space  $BV(Q)$  is endowed with the norm

$$\|u\|_{BV} = \|u\|_{L^1(Q)} + \|\nabla u\|. \quad (6)$$

We say a measurable set  $E \subseteq Q$  has *finite perimeter* in  $Q$  if its indicator function  $\chi_E \in BV(Q)$ . If  $u \in BV(Q)$ , almost all its level sets  $[u \geq \lambda] = \{x \in Q : u(x) \geq \lambda\}$  are sets of finite perimeter (Ambrosio et al., 2000).

Let  $H^1$  denote the one-dimensional Hausdorff measure in  $\mathbb{R}^2$  (a measure of length). If  $E$  is a set of finite perimeter in  $Q$ , the subset of points of the topological boundary of  $E$  where the outer unit normal is defined (in a measure theoretic sense) is called the essential boundary of  $E$  and is denoted by  $\partial^*E$  (Ambrosio et al., 2000). The set  $\partial^*E$  is rectifiable and has finite  $H^1$  measure (Ambrosio et al., 2000).

By combining the assertions in the last two paragraphs, if  $u \in BV(Q)$ , we may define a normal vector  $\theta(x)$  at almost all points (in the  $H^1$  sense) of almost all level sets of  $u$ . This vector field of normals  $\theta$  can be also defined (hence extended to all  $Q$ ) as the Radon-Nikodym derivative of the measure  $\nabla u$  with respect to  $|\nabla u|$ , i.e., it formally satisfies  $\theta \cdot \nabla u = |\nabla u|$  and, also,  $|\theta| \leq 1$  almost everywhere. We shall refer to the vector field  $\theta$  as the vector field of unit normals of the topographic map of  $u$ .

The space of functions of bounded variation  $BV(Q)$  has been frequently used in image processing, mainly in image denoising and restoration problems (Rudin et al., 1992). The space of  $BV$  functions is a reasonable functional model for images since it contains the family of piecewise smooth functions, i.e, functions that are smooth away from its jump discontinuities, which form a set of rectifiable curves. These discontinuities could be identified with edges. Its ability to describe textures is less clear, some textures can be described but to a certain scale of oscillation. An interesting experimental discussion of the adequacy of the  $BV$ -model to describe real images can be seen in Gousseau and Morel (2001). On the other hand, as explained in last paragraph, if  $u$  is a  $BV$  function, then its topographic map can be described by means of the vector field of unit normals of the family of its level lines. For further information concerning functions of bounded variation, we refer to Ambrosio et al. (2000), Evans and Gariepy (1992), and Ziemer (1989).

### 3. The Energy Functional: Continuous and Discrete Descriptions

We first describe the energy functional in a continuous framework, and then describe it at the discrete level.

#### 3.1. The Continuous Formulation

To fix ideas, assume that the multispectral image is given by a function  $\vec{u} : \Omega \rightarrow \mathbb{R}^4$  where  $\Omega$  is a rectangle of  $\mathbb{R}^2$ , say  $[0, 1]^2$ . We shall denote the coordinates of  $\vec{u}$  by  $(X_1, X_2, X_3, X_4)$ . In our experiments below, they represent the red, green, blue and near infrared channels, although they could be different spectral channels. The extension to any set of spectral channels will be immediate. As above, let  $u$  denote the intensity image corresponding to  $\vec{u}$ . We shall assume that the image  $u$  and the spectral channels  $X_n$ ,  $n = 1, 2, 3, 4$ , are aligned.

Assume we are given the image  $u$  on  $\Omega$  and that we know the values of  $\vec{u}$  on a sampling grid  $S \subseteq \Omega$  whose points will be called the low-resolution pixels. Let us denote the known values of  $\vec{u}$  on  $S$  by  $\vec{u}^S = (X_1^S, X_2^S, X_3^S, X_4^S)$ . Recall that the low-resolution pixel is formed from the high-resolution pixel by low-pass filtering followed by subsampling. For each spectral channel  $X_n$ ,  $n \in \{1, 2, 3, 4\}$ , let  $k_n$  be the impulse response of this filter. Our purpose will be to reconstruct  $\vec{u}$  from the data  $u$  and  $\vec{u}^S$ . According to Section 1, we should impose the following two relations:

$$\begin{aligned} u(x) &= \alpha_1 X_1(x) + \alpha_2 X_2(x) + \alpha_3 X_3(x) + \alpha_4 X_4(x), \\ x &= (x_1, x_2) \in \Omega, \end{aligned} \quad (7)$$

where  $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \geq 0$ ,  $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$ , are the coefficients that give the intensity image in terms of the spectral channels, and

$$\begin{aligned} X_n^S(i, j) &= k_n * X_n(i, j), \quad \forall (i, j) \in S, \\ n &= 1, 2, 3, 4, \end{aligned} \quad (8)$$

which correspond to (1) and (2), respectively. We need to assume that it is possible to evaluate  $k_n * X_n$  at any point of  $S$ . For that, we shall assume that  $k$  is the kernel of a convolution operator mapping  $L^2(\Omega)$  into  $C(\overline{\Omega})$  where  $k$  is any of the kernels  $k_1, k_2, k_3, k_4$ . Under this assumption, for any point  $(i, j) \in \Omega$ , the map from any  $f \in L^2(\Omega)$  to

$$k * f(i, j) = \int_{\Omega} k((i, j) - (x_1, x_2)) f(x_1, x_2) dx_1 dx_2$$

is a linear functional in  $L^2(\Omega)$ , and relation (8) makes sense.

The problem of recovering  $(X_1, X_2, X_3, X_4)$  from  $u$  and  $(X_1^S, X_2^S, X_3^S, X_4^S)$  is ill-posed. Indeed, conditions (7) and (8) do not uniquely determine the vector  $(X_1, X_2, X_3, X_4)$ , and the problem involves the inversion of a convolution equation. Typically one is led to a regularization method. The geometric requirement that the geometry of the images  $X_1, X_2, X_3, X_4$  is given by the geometry of the intensity  $u$  will give the required regularization.

*Constraining the Geometry of the Spectral Channels.* We shall assume that the intensity image  $u$  is a function of bounded variation, i.e.,  $u \in BV(\Omega)$ . Since  $\Omega \subseteq \mathbb{R}^2$ , we have in particular that  $u \in L^2(\Omega)$ . As recalled in Section 2, there is a vector field  $\theta : \Omega \rightarrow \mathbb{R}^2$  with  $|\theta| \leq 1$  such that

$$\begin{aligned} \theta \cdot \nabla u &= |\nabla u| \quad \text{as measures in } \Omega, \text{ and} \\ |\theta| &= 1 \quad |\nabla u| \text{ a.e..} \end{aligned}$$

$\theta$  represents the vector field of unit normals of the level lines of  $u$ . Its counterclockwise rotation of angle  $\pi/2$ , denoted by  $\theta^\perp$ , represents the unit tangents of the level lines of  $u$ . Formally, the integration of the system of differential equations  $\dot{x} = \theta^\perp(x)$  would give us again those level lines. Thus, we use either  $\theta$  or  $\theta^\perp$  as a descriptor of the geometry of  $u$  given by its topographic map. In practice, at the discrete level,  $\theta$  can be defined by the relation  $\theta(x) = \frac{\nabla u(x)}{|\nabla u(x)|}$  when  $\nabla u(x) \neq 0$ , and  $\theta(x) = 0$  when  $\nabla u(x) = 0$ ,  $x \in \Omega$ .

Given the vector field  $\theta$ , and assuming that  $X_1, X_2, X_3, X_4 \in BV(\Omega)$ , we impose that the family of lines of  $X_n$ ,  $n = 1, 2, 3, 4$ , is contained in the family of level lines of  $u$  by imposing either that

$$|\nabla X_n| = \theta \cdot \nabla X_n, \quad n = 1, 2, 3, 4, \quad (9)$$

or

$$\theta^\perp \cdot \nabla X_n = 0, \quad n = 1, 2, 3, 4. \quad (10)$$

Both relations mean that  $\theta^\perp$  is tangent to the level lines of  $X_n$ . Thus, heuristically, the level lines of  $X_n$  are integral curves of  $\theta^\perp$ , and, thus, are contained in the family of level lines of  $u$ . We can say that the channels  $X_n$  share the geometry of the panchromatic image. Obviously, relations (9) and (10) cannot be exactly satisfied and we have to impose them in a variational framework (together with the other constraints discussed in the introduction). To help impose (9) we propose including

the following sums in the energy functional

$$\sum_{n=1}^4 \gamma_n \int_{\Omega} (|\nabla X_n| - \theta \cdot \nabla X_n) dx. \quad (11)$$

In a similar way, relation (10) can be imposed in a variational framework by including the following terms in the energy functional

$$\sum_{n=1}^4 \gamma_n \int_{\Omega} |\theta^\perp \cdot \nabla X_n|^2 dx. \quad (12)$$

In both cases, the constants  $\gamma_1, \gamma_2, \gamma_3, \gamma_4 > 0$  permit control of the relative weight assigned to each channel. In practice, we do not privilege any channel over the other ones, so we assign  $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 1$ .

Let us explain the different exponents appearing in the variational formulations of (9) and (10) represented by formulas (11) and (12), respectively. The energy functional constructed with (11), plus the terms corresponding to constraints (7) and (8), can be minimized in  $BV(\Omega)^4 = BV(\Omega) \times BV(\Omega) \times BV(\Omega) \times BV(\Omega)$  (see Theorem 1 in Appendix 5). As explained in Section 2, we use this space because we want to ensure the possibility that the solutions of our problem have discontinuities along rectifiable curves. The energy functional constructed with (12) can be minimized in the fourth power of the space of functions  $X \in L^2(\Omega)$  such that  $\theta^\perp \cdot \nabla X \in L^2(\Omega)$  (see Appendix 5), and this set of functions also permits to recover the discontinuities in the images, as long as  $\theta$  is orthogonal to them. The expression (12) only measures the regularity of  $X_n$  along the integral curves of  $\theta^\perp$ , which are the level lines of  $u$ , and the edges of the panchromatic image are not penalized. To measure such regularity, we could use a different power  $p > 1$  instead of  $p = 2$  (the case  $p = 1$  can also be used in practice, though at the theoretical level it introduces some mathematical difficulties that we cannot handle at this moment); we have tested it without observing noticeable differences. For computational reasons, we have chosen  $p = 2$  since the corresponding Euler-Lagrange equations are linear in the unknowns and the numerical algorithm to solve them is faster. We also notice that (12) is invariant under contrast inversion. Such invariance is particularly suitable when considering the case of false color images (see (Lisani and Morel, 2000)). In any case, we have tested numerically both functionals with very similar results (see Section 4.2).

Let us write functional (11) in a convenient form for computational purposes. For that, we integrate the second term of each integral by parts. By using  $\theta \cdot \nu = 0$  on  $\partial\Omega$  (where  $\nu$  is the outer unit normal to  $\partial\Omega$ ), we may write (11) in the form

$$\sum_{n=1}^4 \gamma_n \int_{\Omega} (|\nabla X_n| + \operatorname{div} \theta \cdot X_n) dx. \quad (13)$$

Recall that we should not use the symbol  $dx$  when we integrate a measure.

*Imposing (7) and (8) in a Variational Framework.* We impose the constraint (7) by minimizing the integral term

$$\int_{\Omega} (\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 - u)^2 dx. \quad (14)$$

To help impose constraint (8), we include the following sums in the objective function

$$\sum_{(i,j) \in S} (k_n * X_n(i,j) - X_n^S(i,j))^2 \quad (15)$$

for each  $n = 1, 2, 3, 4$ . We rewrite the above relations in an integral form; this will be useful for writing the Euler-Lagrange equations in a more compact way. We need some notation for that. Let  $\delta_{(i,j)}$  be Dirac's delta function located at the point  $(i, j)$ . Let  $\Pi_S = \sum_{(i,j) \in S} \delta_{(i,j)}$  be a Dirac's comb defined by the grid  $S$ . Then, we may write (15) in integral form as

$$\int_{\Omega} \Pi_S (k * X_n(x) - X_n^S(x))^2 dx \quad (16)$$

for each  $n = 1, 2, 3, 4$ , where  $X_n^S(x)$  denotes an arbitrary extension of  $X_n^S(i, j)$  as a continuous function from  $S$  to  $\Omega$ . Since the integrand term above is multiplied by  $\Pi_S$ , the integral expression (16) does not depend on the particular extension of  $X_n^S$ ,  $n = 1, 2, 3, 4$ .

We shall impose a further constraint on  $X_1, X_2, X_3, X_4$ . Indeed, for each  $n = 1, 2, 3, 4$ , let

$$M_n = \max_{(i,j) \in S} \max \left( \frac{u(i,j)}{\alpha_n}, X_n^S(i,j) \right). \quad (17)$$

Then, we shall impose that

$$0 \leq X_n \leq M_n, \quad n = 1, 2, 3, 4. \quad (18)$$

These constraints are useful for mathematical justification of the algorithm.

*The Energy Functional.* To fix ideas, we propose computing the high-resolution multispectral images  $X_1, X_2, X_3, X_4$  by minimizing the energy functional:

$$\begin{aligned} & \sum_{n=1}^4 \gamma_n \int_{\Omega} |\theta^\perp \cdot \nabla X_n|^2 dx \\ & + \lambda \int_{\Omega} \left( \sum_{n=1}^4 \alpha_n X_n - u \right)^2 dx \\ & + \mu \sum_{n=1}^4 \int_{\Omega} \Pi_S ((k_n * X_n(x) - X_n^S(x))^2 dx \end{aligned} \quad (19)$$

subject to  $0 \leq X_n \leq M_n$ ,  $n = 1, 2, 3, 4$ , where  $X_n \in BV(\Omega)$ ,  $\gamma_n, \lambda, \mu > 0$ ,  $n = 1, 2, 3, 4$  (in practice, all these parameters are taken equal to 1). As an alternative based on (13), we could use the functional

$$\begin{aligned} & \sum_{n=1}^4 \gamma_n \int_{\Omega} (|\nabla X_n| + \operatorname{div} \theta \cdot \nabla X_n) dx \\ & + \lambda \int_{\Omega} \left( \sum_{n=1}^4 \alpha_n X_n - u \right)^2 dx \\ & + \mu \sum_{n=1}^4 \int_{\Omega} \Pi_S ((k_n * X_n(x) - X_n^S(x))^2 dx \end{aligned} \quad (20)$$

subject to the same constraints as (19). We notice that the first integral in (20) was also used in Ballester et al. (2001) to impose the constraint  $\theta \cdot \nabla u = |\nabla u|$  in the context of a variational formulation of filling-in holes in images by joint interpolation of vector fields and gray levels. In that context, both  $\theta$  and  $u$  were unknown.

If we consider only the channels  $X_1 = \text{red}$ ,  $X_2 = \text{green}$ ,  $X_3 = \text{blue}$ , both functionals (19) and (20) can be effectively used as we have shown in Ballester et al. (2003). When we add the channel  $X_4 = \text{near infrared}$ , it seems slightly better to use functional (19), or a variant of it that uses (10) for the near infrared channel and (9) for the red, green, and blue channels. The main reason is that functional (19) is invariant under the change  $\theta \rightarrow -\theta$ , or, in other words, it is invariant under contrast inversion. The main advantage of (19) is that the corresponding Euler-Lagrange equations are linear in  $X_n$  and a steepest descent algorithm converges to a minimum faster than in the case of (20). Indeed, for

an image of size  $800 \times 800$  the execution time is 63 seconds on a 1.8 GHz Pentium.

### 3.2. The Discrete Formulation

To proceed with the discrete numerical algorithm, we assume that the panchromatic image  $u$  is given on  $\{0, 1, \dots, N-1\} \times \{0, 1, \dots, N-1\}$ . We replace the gradients in (20) by discrete approximations; for any scalar function  $f$  we shall use the notation

$$\begin{aligned}\nabla^{+,+} f &= (\nabla_x^+ f, \nabla_y^+ f), & \nabla^{+,-} f &= (\nabla_x^+ f, \nabla_y^- f), \\ \nabla^{-,+} f &= (\nabla_x^- f, \nabla_y^+ f), & \nabla^{-,-} f &= (\nabla_x^- f, \nabla_y^- f)\end{aligned}$$

where

$$\begin{aligned}\nabla_x^+ f(i, j) &= f(i+1, j) - f(i, j), \\ \nabla_x^- f(i, j) &= f(i, j) - f(i-1, j), \\ \nabla_y^+ f(i, j) &= f(i, j+1) - f(i, j), \\ \nabla_y^- f(i, j) &= f(i, j) - f(i, j-1).\end{aligned}$$

Note that the dual operators to  $\nabla^{+,+}$ ,  $\nabla^{+,-}$ ,  $\nabla^{-,+}$ ,  $\nabla^{-,-}$  are, respectively, the operators  $div^{-,-}$ ,  $div^{-,+}$ ,  $div^{+,-}$ ,  $div^{+,+}$ . We use the notation  $\theta_{\alpha,\beta} = \frac{\nabla^{\alpha,\beta} u}{|\nabla^{\alpha,\beta} u|}$  if  $\nabla^{\alpha,\beta} u \neq 0$  and  $\theta_{\alpha,\beta} = 0$  if  $\nabla^{\alpha,\beta} u = 0$ , for  $\alpha, \beta = +, -$ . We observe that in our experiments in Section 4.2 the values of  $u$  and  $X_n$ ,  $n = 1, 2, 3, 4$ , which were furnished to us by the CNES, are quantized in the range  $[0, 511]$ , or  $[0, 1023]$ , depending on the set of images. To dequantize the data and have a better estimate of  $\theta$  we can apply a dequantization strategy, as the one proposed in Desolneux et al. (2002). We have indeed tested this procedure with the CNES data without significant improvement of our results.

For brevity, we shall only describe the discrete formulation of (19), which can be written as

$$\begin{aligned}& \sum_{n=1}^4 \frac{\gamma_n}{4} \sum_{\alpha,\beta=+,-} \sum_{i,j=0}^{N-1} |\theta_{\alpha,\beta}(i, j)^\perp \cdot \nabla^{\alpha,\beta} X_n(i, j)|^2 \\ & + \lambda \sum_{i,j=0}^{N-1} \left( \sum_{n=1}^4 \alpha_n X_n(i, j) - u(i, j) \right)^2 \\ & + \mu \sum_{n=1}^4 \sum_{(i,j) \in S} ((k_n * X_n(i, j) - X_n^S(i, j))^2),\end{aligned}$$

subject to

$$0 \leq X_n \leq M_n, \quad n = 1, 2, 3, 4. \quad (21)$$

Observe that we have simultaneously used the four discretizations for the gradient, since using only one of the approximations may produce some artifacts or asymmetries in the results.

## 4. Algorithm and Numerical Experiments

To be concise, we only describe the algorithm used to minimize (21). Applying the gradient descent method, to minimize (21) we iteratively compute the solution using the equations

$$\begin{aligned}X_n^{(m+1)} &= X_n^{(m)} + \Delta t \frac{\gamma_n}{4} \\ & \sum_{\alpha,\beta=+,-} \operatorname{div}^{\alpha,\beta*} \left( (\theta_{\alpha,\beta}^\perp \cdot \nabla^{\alpha,\beta} X_n^{(m)}) \theta_{\alpha,\beta}^\perp \right) \\ & - \mu \Delta t k_n^t * (\Pi_S(k_n * X_n^{(m)} - X_n^S)) \\ & - \lambda \alpha_n \Delta t \left( \sum_{n=1}^3 \alpha_n X_n^{(m)} - u \right).\end{aligned} \quad (22)$$

where  $n = 1, 2, 3, 4$ ,  $m \geq 0$ . The constraint that  $X_n$ ,  $n = 1, 2, 3, 4$ , should remain in the range between 0 and an upper value  $M_n$  can be imposed after each iteration by truncation. To avoid a cumbersome expression we have avoided writing in detail the terms involving  $k_n$ . In practice, it is very important to guarantee that the energy (21), call it  $E(X_1^{(m)}, X_2^{(m)}, X_3^{(m)}, X_4^{(m)})$ , decreases along the evolution, i.e., that

$$\begin{aligned}E(X_1^{(m+1)}, X_2^{(m+1)}, X_3^{(m+1)}, X_4^{(m+1)}) \\ \leq E(X_1^{(m)}, X_2^{(m)}, X_3^{(m)}, X_4^{(m)}).\end{aligned} \quad (23)$$

For that, we have to control the time increment  $\Delta t$ . Indeed, at each iteration, we only accept  $\Delta t$  if (23) holds. We have the possibility of choosing  $\Delta t$  so that the energy has the largest decrease in the direction of the energy gradient. If  $\nabla E(X_1^{(m)}, X_2^{(m)}, X_3^{(m)}, X_4^{(m)})$  denotes the energy gradient given in the right-hand side of (22), then  $\Delta t$  can be chosen as a solution of

$$\begin{aligned}\min_{s>0} E((X_1^{(m)}, X_2^{(m)}, X_3^{(m)}, X_4^{(m)}) \\ - s \nabla E(X_1^{(m)}, X_2^{(m)}, X_3^{(m)}, X_4^{(m)})).\end{aligned} \quad (24)$$

In practice, we observed that long-term decrease of the energy is favored not with the optimal choice of  $\Delta t$  at each iteration but with a constant choice, adapted so that (23) holds at each iteration.

#### 4.1. Description of the Data

We test our method with some reference multispectral satellite images furnished to us by the CNES. The data consists of a panchromatic image  $u(i, j)$ ,  $i, j \in \{0, \dots, N - 1\}$ , at a resolution of 0.7 m/pixel, and a set of multispectral images

$$\vec{u}^{2.8} = (R_{2.8}, G_{2.8}, B_{2.8}, IR_{2.8})$$

at a resolution of 2.8 m/pixel, which represent the red, green, blue, and near infrared spectral channels, respectively. These data have been generated by the CNES simulating a precise satellite. To test our experiments, we also employ a ‘ground truth’ reference image

$$\vec{u}^{\text{ref}} = (R_{\text{ref}}(i, j), G_{\text{ref}}(i, j), B_{\text{ref}}(i, j), IR_{\text{ref}}(i, j)), \\ i, j \in \{0, \dots, N - 1\},$$

at a resolution of 0.7 m/pixel.

Following the notation used above, we identify the multispectral channels as

$$X_1^S = R_{2.8}, X_2^S = G_{2.8}, X_3^S = B_{2.8}, X_4^S = IR_{2.8},$$

and, similarly, our unknowns  $(X_1, X_2, X_3, X_4)$  represent the red, green, blue, and near infrared channels, respectively.

Each satellite acquires a different panchromatic image, hence it is described by a different set of values  $\alpha_i$ ,  $i = 1, 2, 3, 4$ . For an account of different satellite models we refer to <http://smssc.cnes.fr>. For reasons of space, we have restricted our experiments below to the simulation of a precise case, where the panchromatic constraint (7) is given by

$$u = 0.5X_1 + 0.5X_2. \quad (25)$$

In other words, the constants  $\alpha_n$  are

$$\alpha_1 = 0.5, \quad \alpha_2 = 0.5, \quad \alpha_3 = 0, \quad \alpha_4 = 0.$$

The impulse responses  $k_n$  have also been provided to us by the CNES, and have a size of  $41 \times 41$ . Both  $\alpha_n$ , and  $k_n$ ,  $n = 1, 2, 3, 4$ , are the parameters used to describe the image acquisition model.

#### 4.2. Numerical Experiments

The experiments below test an increase of resolution by a factor of 4. We display our results in the case of true color (where we display the red, green, and blue channels) and in the case of false color (where we display the near infrared, red, and green channels).

We iteratively minimize (21) using the gradient descent equations given in (22). We initialize (22) with an image obtained by simple replication of  $\vec{u}^{2.8}$  by a factor 4. We have also tested other initializations like a DCT zoom (of factor 4), or a more sophisticated initialization obtained using the approach in Caselles et al. (2000), and they speed up the convergence of the algorithm. But our experiments have also shown that the results obtained are similar.

The functions  $u$  and  $\vec{u}^{2.8}$  constitute the basic data of the problem. We stress that the unknowns are the four spectral channels  $X_n$ ,  $n = 1, 2, 3, 4$ , representing the red, green, blue, and near infrared channels. But, for purposes of visualization, we first display true color images (involving the  $(R, G, B)$  channels), and then false color images (involving the  $(IR, R, G)$  channels).

The reference image permits us to assess the quality of the reconstructed image. We shall compare the reconstructed images  $(X_1, X_2, X_3, X_4) = (R, G, B, IR)$  with the reference images  $(R_{\text{ref}}, G_{\text{ref}}, B_{\text{ref}}, IR_{\text{ref}})$  both visually and by displaying some error measures. We



Figure 1. Experiment 1: The true color reference image  $(R_{\text{ref}}, G_{\text{ref}}, B_{\text{ref}})$ .

shall compute the errors on the whole image and on certain representative regions (see Fig. 8) . Let  $A$  denote the region where the error is computed (the whole image or a region of it). We shall use the maximum,  $\ell^1$ , and  $\ell^2$  error measures, i.e.,

(i) The maximum error:

$$\max_{(i,j) \in A} |R(i, j) - R_{\text{ref}}(i, j)|.$$



Figure 2. Experiment 1: (a) Left: The panchromatic image at resolution 0.7 m/pixel, denoted in the text by  $u$ . (b) Right:  $(R_{2.8}, G_{2.8}, B_{2.8})$  spectral channels at a resolution of 2.8 m/pixel, i.e., the  $(R, G, B)$  coordinates of  $\vec{u}^{2.8}$ .

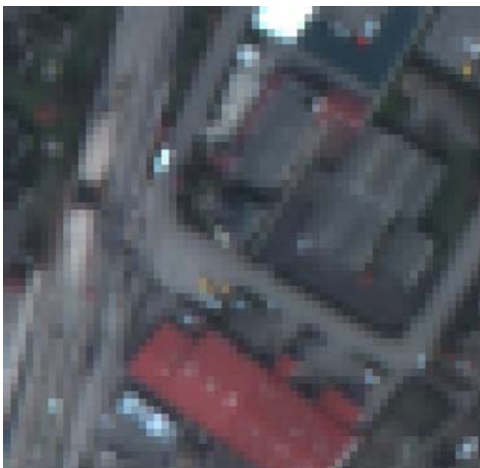


Figure 3. Experiment 1:  $(R, G, B)$  channels of the initialization (obtained by replication by a factor of four) of the algorithm based on functional (21).

(ii) The  $\ell^1$  norm

$$\frac{1}{|A|} \sum_{(i,j) \in A} |R(i, j) - R_{\text{ref}}(i, j)|$$

where  $|A|$  denotes the number of pixels of  $X$ .

(iii) The  $\ell^2$  norm

$$\left( \frac{1}{|A|} \sum_{(i,j) \in A} |R(i, j) - R_{\text{ref}}(i, j)|^2 \right)^{1/2},$$



Figure 4. Experiment 1:  $(R, G, B)$  channels of the reconstructed image (at the resolution of the panchromatic image) obtained using functional (21).



Figure 5. Experiment 1: False color image showing the  $(I_{R_{\text{ref}}}, R_{\text{ref}}, G_{\text{ref}})$  channels of the reference image  $\vec{u}^{\text{ref}}$ .



Figure 6. Experiment 1:  $(IR_{2.8}, R_{2.8}, G_{2.8})$  spectral channels at resolution 2.8 m/pixel, i.e., the  $(IR, R, G)$  channels of  $\bar{u}^{2.8}$ .



Figure 7. Experiment 1:  $(IR, R, G)$  channels of the reconstructed image (at the resolution of the panchromatic image) obtained using functional (21).

with similar expressions for the other channels. We also display histograms of errors, and compute some explicit percentiles.

We performed experiments on a set of 31 images furnished to us by the CNES who also evaluated the results. They were good for the  $(R, G, B)$  channels and of poorer quality for the infrared one. This is not a surprise, since our basic assumption that the geometry of the spectral channels is given by the geometry of the panchromatic image is not so reasonable for the infrared one. For reasons of space we present only two experiments with its corresponding errors. In Experiment 1, the range of values of  $u$  and  $X_n$  is the interval  $[0, 511]$ . In experiment 2, this range is  $[0, 1023]$ .

*Experiment 1: Display of True Color Results.* Figure 1 displays the red, green, and blue channels of reference image  $\bar{u}^{\text{ref}}$ , i.e.,  $(R_{\text{ref}}, G_{\text{ref}}, B_{\text{ref}})$ . Figure 2(a) displays the panchromatic image  $u$ , and Fig. 2(b) shows the spectral channels  $(R_{2.8}, G_{2.8}, B_{2.8})$  of the lower resolution data  $\bar{u}^{2.8}$ . Both images were furnished to us by the CNES.

Figure 4 displays the  $(R, G, B)$  channels of the reconstructed image, which are at the resolution of the panchromatic image, obtained using functional (21). The  $k_n * X_n$  convolution was computed in the Fourier

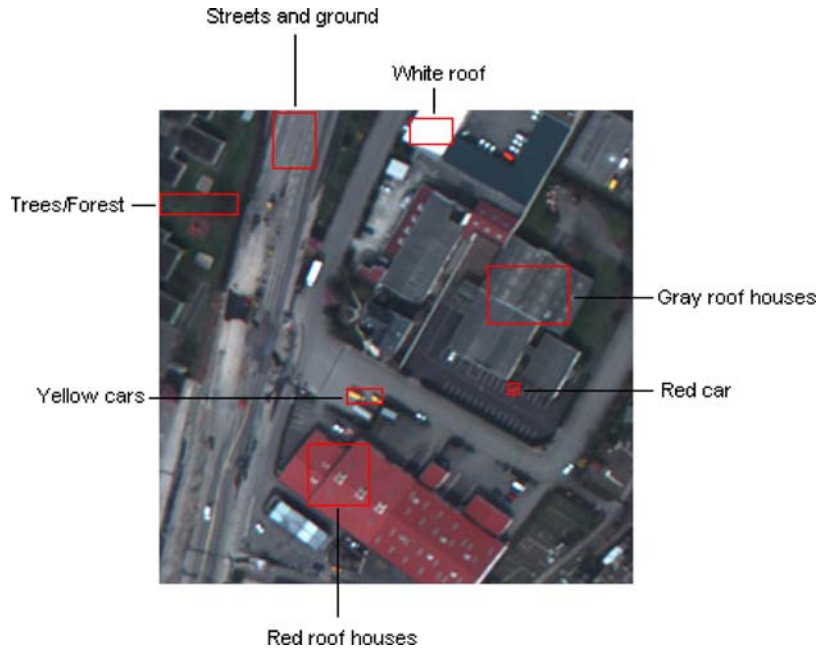


Figure 8. Experiment 1: Regions where the statistics of Fig. 9 are computed.

	Errors									
	Max	L <sup>1</sup>	L <sup>2</sup>	Mode	P <sub>10</sub>	P <sub>25</sub>	P <sub>50</sub>	P <sub>75</sub>	P <sub>90</sub>	
<b>B0 - Blue</b>	129.9	5.4	8.3	1	0.6	1.5	3.5	6.9	12.0	
Trees/forest	14.4	2.8	3.5	1	0.6	1.2	2.4	3.9	5.8	
Red roof houses	33.4	4.1	5.9	1	0.6	1.3	2.8	5.6	9.6	
Gray roof houses	26.8	3.8	5.4	1	0.5	1.3	2.8	5.1	8.3	
White roof	59.4	10.5	15.6	2	1.2	2.7	6.1	14.1	26.7	
Streets and ground	26.3	6.0	7.5	1	0.9	2.4	5.0	8.3	12.8	
Red car	26.0	8.3	10.0	8	2.3	4.4	7.2	11.6	13.4	
Yellow cars	29.0	7.4	9.8	1	1.0	2.2	5.2	10.4	18.0	
<b>B1 - Green</b>	64.2	3.3	5.0	1	0.4	1.0	2.3	4.3	7.2	
Trees/forest	11.7	1.8	2.5	1	0.2	0.6	1.4	2.5	4.1	
Red roof houses	30.5	3.6	5.3	1	0.4	1.0	2.4	4.7	8.3	
Gray roof houses	19.8	2.4	3.2	1	0.4	0.9	1.9	3.3	5.1	
White roof	23.5	4.3	5.9	1	0.6	1.6	3.3	5.4	8.6	
Streets and ground	21.4	2.7	3.8	1	0.4	1.0	2.1	3.7	5.8	
Red car	60.2	16.0	21.0	6	4.2	6.1	12.1	22.0	30.1	
Yellow cars	37.7	8.1	11.6	1	0.9	1.7	5.5	11.7	18.2	
<b>B2 - Red</b>	67.9	3.2	4.9	1	0.4	1.0	2.2	4.1	6.9	
Trees/forest	8.0	1.9	2.4	1	0.3	0.8	1.5	2.8	4.0	
Red roof houses	33.9	3.6	5.4	1	0.4	1.1	2.5	4.8	8.2	
Gray roof houses	15.9	2.3	3.1	1	0.3	0.8	1.7	3.1	4.9	
White roof	26.0	4.6	5.8	1	0.8	1.8	3.8	6.5	8.7	
Streets and ground	22.4	2.9	4.0	1	0.4	1.0	2.2	4.0	6.1	
Red car	54.6	13.4	18.3	1	1.2	5.5	11.1	17.1	28.2	
Yellow cars	44.4	7.6	11.3	2	0.7	1.9	4.4	10.2	17.8	
<b>B3 - Near Infrared</b>	249.2	14.6	22.5	1	1.3	3.6	8.8	19.2	35.3	
Trees/forest	88.4	13.5	18.1	6	2.0	5.1	10.0	19.1	29.3	
Red roof houses	70.0	8.9	12.9	1	1.0	2.6	6.2	11.8	20.1	
Gray roof houses	75.4	5.3	7.6	3	0.8	2.0	4.0	6.9	10.7	
White roof	77.8	18.8	26.5	2	2.2	4.4	11.8	25.1	52.1	
Streets and ground	111.3	11.9	16.8	4	1.8	4.7	9.3	15.3	23.7	
Red car	81.3	44.1	48.3	31	12.1	31.7	44.7	58.8	67.0	
Yellow cars	119.9	29.3	38.9	2	2.6	9.1	22.5	43.6	62.5	

Histogram

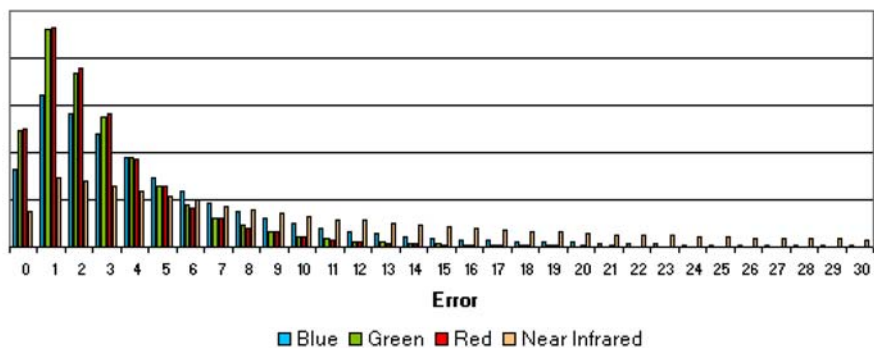


Figure 9. Errors corresponding to Experiment 1: Columns 1,2,3 contain the maximum, the  $\ell^1$  and the  $\ell^2$  errors, respectively. Column 4 contains the mode of the error distribution, that is, the most probable error. Columns 5 to 9 contain the percentiles of the error, i.e.,  $P_\alpha$  is the value such that  $P(\text{error} \leq P_\alpha) = \alpha/100$ .

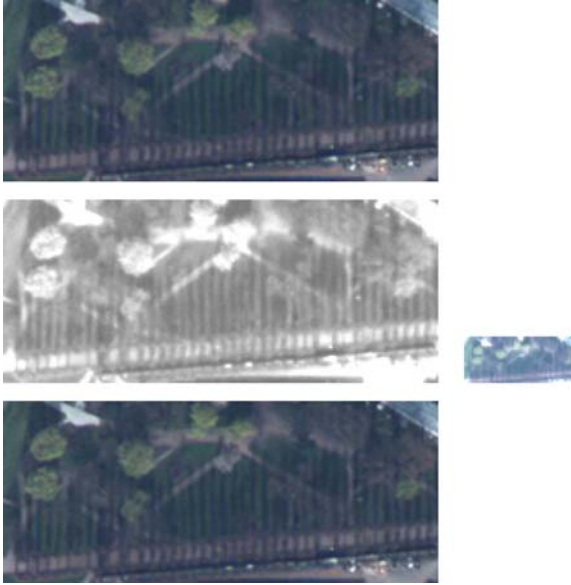


Figure 10. Experiment 2: (a) Top: The true color reference image ( $R_{\text{ref}}, G_{\text{ref}}, B_{\text{ref}}$ ). (b) Middle left: The panchromatic image at resolution 0.7 m/pixel. (c) Middle right: ( $R_{2.8}, G_{2.8}, B_{2.8}$ ) spectral channels at a resolution of 2.8 m/pixel. (d) Bottom: ( $R, G, B$ ) channels of the reconstructed image (at the resolution of the panchromatic image) obtained using functional (21).

domain. The result should be compared with the reference image in Fig. 1. The ( $R, G, B$ ) channels used to initialize (22) are displayed in Fig. 3.

Figure 9 displays the errors between the reference image and the result; in particular, it contains the errors in the ( $R, G, B$ ) channels. Figure 8 displays some regions where errors are computed. Below the table, we display the corresponding histogram of errors.

*Experiment 1: Display of False Color Results.* To display a false color image we have used the near infrared, red and green channels in place of the usual ( $R, G, B$ ) channels (and in this order). Figure 5 displays the ( $IR_{\text{ref}}, R_{\text{ref}}, G_{\text{ref}}$ ) channels of the reference image  $\vec{u}^{\text{ref}}$ . Figure 6 displays the spectral channels ( $IR_{2.8}, R_{2.8}, G_{2.8}$ ) at a resolution of 2.8 m/pixel. Figure 7 displays the ( $IR, R, G$ ) channels of the reconstructed image, which are at the resolution of the panchromatic image, obtained using functional (21).

Figure 9 displays the errors between the reference image and the result; in particular, it contains the errors

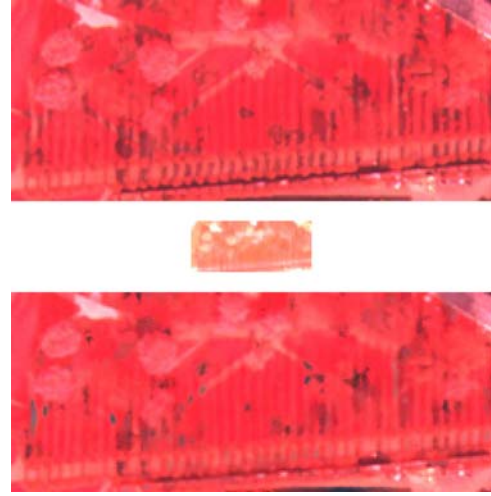


Figure 11. Experiment 2: (a) Top: False color image of the ( $IR_{\text{ref}}, R_{\text{ref}}, G_{\text{ref}}$ ) channels of the reference image. (b) Middle: ( $IR_{2.8}, R_{2.8}, G_{2.8}$ ) spectral channels at a resolution of 2.8 m/pixel. (c) Bottom: ( $IR, R, G$ ) channels of the reconstructed image (at the resolution of the panchromatic image) obtained using functional (21).

in the IR channel. Below the table, we display the corresponding histogram of errors.

*Experiment 2: True Color Results.* Figure 10 displays an experiment with an image of vegetation. Figure 10(a) displays the ( $R, G, B$ ) channels of the reference image. Figure 10(b) displays the panchromatic image at a resolution of 0.7 m/pixel, and Fig. 10(c) displays the ( $R, G, B$ ) channels at the resolution of 2.8 m/pixel. Figure 10(d) displays the ( $R, G, B$ ) channels of the reconstructed image at the resolution of 0.7 m/pixel obtained using functional (21). The corresponding errors are displayed in Fig. 12.

*Experiment 2: False Color Results.* Figure 11(a) displays the ( $IR, R, G$ ) channels of the reference image. Figure 11(b) displays the ( $IR, R, G$ ) channels at the resolution of 2.8 m/pixel. Figure 11(c) displays the reconstructed ( $IR, R, G$ ) channels at the resolution of 0.7 m/pixel obtained using functional (21). The corresponding errors are displayed in Fig. 12.

*Remark.* We have also tested functional (20) and functional (19) replacing the power  $p = 2$  in the

	Errors								
	Max	$L^1$	$L^2$	Mode	$P_{10}$	$P_{25}$	$P_{50}$	$P_{75}$	$P_{90}$
<b>B0 - Blue</b>	93.3	5.4	7.7	1	0.7	1.8	4.0	7.2	11.4
<b>B1 - Green</b>	65.5	4.5	6.2	1	0.6	1.6	3.5	6.1	9.4
<b>B2 - Red</b>	65.9	4.7	6.4	1	0.7	1.7	3.7	6.5	9.8
<b>B3 - Near Infrared</b>	593.6	75.1	102.8	1	9.5	24.5	55.0	103.6	167.3

Figure 12. Errors corresponding to Experiment 2. Columns 1,2,3 contain the maximum, the  $\ell^1$  and the  $\ell^2$  errors, respectively. Column 4 contains the mode of the error distribution. Columns 5 to 9 contain the percentiles of the error.



Figure 13. (a) Top:  $(R, G, B)$  channels of the reconstructed image obtained by interpolating the chromatic channels in the CIELAB color system. (b) Bottom:  $(IR, R, G)$  channels of the reconstructed image with the same method.

term  $\int_{\Omega} |\theta^{\perp} \cdot \nabla X_n|^p dx$  by  $p = 1$  with comparable results. The results obtained with functional (19) with  $p = 2$  are usually better than the ones with (20) and similar to the ones obtained with  $p = 1$ , but, in any case, the errors differ in the last digit by a small amount.

*Interpolation of Chromatic Channels in the CIELAB Color System.* We display an experiment to compare our results with a simple fusion method based on the following strategy: (A) We transform the  $(R, G, B)$  channels of the low-resolution multispectral image to the  $L^*a^*b^*$  color system (see (Hunt, 1998)), (B) We interpolate the  $a^*, b^*$  chromatic components at the resolution of the panchromatic image using bilinear or bicubic interpolation, (C) Using the panchromatic image as the  $Y$  component in the  $XYZ$  color system, we compute the corresponding  $L^*$  component. Then, together with the  $a^*, b^*$  channels obtained in Step B), we compute the  $(R, G, B)$  components of the reconstructed high-resolution image. Even if it is not adapted to this case, we use the same strategy replacing the  $(R, G, B)$  channels by the  $(IR, R, G)$  ones. Notice that this method also imposes the geometry of the panchromatic image to the reconstructed high-resolution multispectral channels.

We perform our experiment on the panchromatic image and the  $(R, G, B)$  spectral channels displayed in Fig. 2. Then we repeat it with the  $(IR, R, G)$  channels displayed in Fig. 6. Figure 13(a) and (b) display the reconstructed  $(R, G, B)$ , respectively  $(IR, R, G)$ , channels obtained using the simple interpolation method described above. The corresponding errors are displayed in Fig. 14. We observe that the errors are slightly larger than the ones obtained with our method (see Fig. 9).

Using a simple method, like bilinear, bicubic, zero padding using the FFT, or DCT, to interpolate the multispectral channels does not give acceptable results.

	Errors								
	Max	L <sup>1</sup>	L <sup>2</sup>	Mode	P <sub>10</sub>	P <sub>25</sub>	P <sub>50</sub>	P <sub>75</sub>	P <sub>90</sub>
<b>B0 - Blue</b>	173.4	6.9	10.6	1	0.8	2.1	4.6	8.9	14.7
<b>B1 - Green</b>	91.5	4.5	6.5	1	0.6	1.5	3.2	5.9	9.5
<b>B2 - Red</b>	77.7	3.7	5.8	1	0.5	1.2	2.5	4.6	7.8
<b>B3 - Near Infrared</b>	240.9	17.6	30.3	1	1.3	3.3	8.2	21.0	46.0

Figure 14. Errors corresponding to the results in Fig. 13. Columns 1,2,3 contain the maximum, the  $\ell^1$  and the  $\ell^2$  errors, respectively. Column 4 contains the mode of the error distribution. Columns 5 to 9 contain the percentiles of the error.

## 5. Conclusions

We have reported a variational model for increasing the resolution of satellite multispectral data knowing the panchromatic image at higher resolution and the multispectral data at a lower resolution. The model incorporates the relations between the spectral channels and the panchromatic image (1) and the relation describing how low-resolution pixels are formed from high-resolution pixels by a lowpass filter followed by subsampling (2). The main feature of our model is the incorporation of the hypothesis that for satellite multispectral images, to a large extent, the geometry of the spectral channels is contained in the topographic map of its panchromatic image. We have constructed two slightly different energy functionals, (20) and (19), which incorporate the above three basic postulates. Note that functional (19) is invariant under contrast inversion, i.e., under the change of  $\theta$  into  $-\theta$ . Contrast inversion seems to be important because of the presence of the near infrared channel. Finally, we described our algorithm to minimize these functionals and, for simplicity, we displayed some experiments relative to functional (19) showing its capabilities in true color and false color. The results with both functionals are comparable.

Finally, let us mention that the Euler-Lagrange equations of functional (19) are linear in the variables and the gradient descent exhibits fast convergence. The number of operations per pixel is 2700. The time spent reconstructing an image of size  $800 \times 800$  pixels on a 1.8 GHz Pentium is 63 seconds.

## Appendix

*Existence of Solutions of the Variational Problems (19) and (20).* Let  $W^{1,2}(\Omega)$  denote the space of functions  $u \in L^2(\Omega)$  such that  $\nabla u \in L^2(\Omega)$ . Assume that  $\theta : \Omega \rightarrow \mathbb{R}^2$  is such that  $|\theta(x)| \leq 1$  a.e., and satisfies  $\operatorname{div} \theta^\perp \in L^2(\Omega)$ . Let  $W(\Omega, \theta)$  be the completion of

$W^{1,2}(\Omega)$  with respect to the norm

$$\Phi(u) = \left( \int_{\Omega} |\theta^\perp \cdot \nabla u|^2 dx \right)^{1/2} + \left( \int_{\Omega} |u|^2 \right)^{1/2}.$$

We have the following result.

**Theorem 1.** *If  $\operatorname{div} \theta^\perp \in L^2(\Omega)$ , then the functional (19) admits a minimum in  $W(\Omega, \theta)^4$ . Similarly, if  $\operatorname{div} \theta \in L^2(\Omega)$ , functional (20) admits a minimum in  $BV(\Omega)^4$ .*

## Acknowledgment

We thank the reviewers for many corrections that improved the paper. In particular, we thank one of them for having suggested the experiment using the  $L^*a^*b^*$  color system. This work was supported by the CNES. L. Igual also acknowledges support by the company THALES (France). C. Ballester, V. Caselles, and J. Verdera acknowledge partial support by the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya and by PNPGC project, reference BFM2003-02125.

## References

- Aiazzi, B., Alparone, L., Baronti, S., and Garzelli, A. 2002. Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Trans. on Geoscience and Remote Sensing*, 40(10):2300–2312.
- Ambrosio, L., Caselles, V., Masnou, S., and Morel, J. 2001. The connected components of sets of finite perimeter. *European Journal of Math*, 3:39–92.
- Ambrosio, L., Fusco, N., and Pallara, D. 2000. Functions of Bounded Variation and Free Discontinuity Problems, *Oxford Mathematical Monographs*.
- Ballester, C., Bertalmio, M., Caselles, V., Sapiro, G., and Verdera, J. 2001. Filling-in by joint interpolation of vector fields and gray levels. *IEEE Transactions on Image Processing*, 10(8):1200–1211.

- Ballester, C., Caselles, V., Rougé, B., and Verdera, J. 2003. Une méthode géométrique de fusion des images P+XS. *Rapport CNES*.
- Caselles, V., Coll, B., and Morel, J. 1999. Topographic maps and local contrast changes in natural images. *Int. J. Comp. Vision*, 33(1):5–27.
- Caselles, V., Coll, B., and Morel, J. 2002. Geometry and color in natural images. *J. Math. Imaging and Vision*, 16:89–107.
- Caselles, V., Sapiro, G., and Chung, D. 2000. Vector median filters, vector morphology, and coupled PDE's: Theoretical connections. *J. Math. Imaging and Vision*, 12:109–120.
- Desolneux, A., Moisan, L., and Morel, J. 2002. Dequantizing image orientation. *IEEE Transactions on Image Processing*, 11:1129–1140.
- Evans, L. and Gariepy, R. 1992. *Measure Theory and Fine Properties of Functions*, Studies in Advanced Math., CRC Press.
- Gousseau, Y. and Morel, J. 2001. Are natural images of bounded variation ? *SIAM J. Math. Anal.*, 33(2001):634–648.
- Guichard, F. and Morel, J. 2003. *Partial differential equations and image iterative filtering*, Forthcoming book.
- Hunt, R. 1998. *Measuring Color*, England: Fountain Press,
- Lisani, J. and Morel, J. 2000. Comparaison morphologique d'images. *Rapport CNES*.
- Monasse, P. 1999. Contrast invariant image registration. *Proc. of the International Conf. on Acoustics, Speech and Signal Processing, Phoenix, Arizona*, 6:3221–3224.
- Rudin, L., Osher, S., and Fatemi, E. 1992. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1992):259–269.
- Serra, J. 1982. *Image analysis and mathematical morphology*. Academic Press.
- Wald, L., Ranchin, T., and Mangolini, M. 1997. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogramm. Eng. Remote. Sensing*, 63(6):691–699.
- Wertheimer, M. 1923. Untersuchungen zur Lehre der Gestalt, II. *Psychologische Forschung*, 4:301–350.
- Ziemer, W. 1989. *Weakly Differentiable Functions*, GTM 120, Springer Verlag.