

# A Variational Model for P+XS Image Fusion

C. Ballester, V. Caselles, J. Verdera  
University Pompeu Fabra  
Departament de Tecnologia  
Pg. Circumvalació 8, 08003 Barcelona, Spain  
*e-mail: vicent.caselles@upf.edu*

B. Rougé  
CNES  
Av. Ed Belin, 31401 Toulouse, France  
*e-mail: Bernard.Rouge@cnes.fr*

## 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 relating 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 to construct an energy functional (with several variants) whose minima will give the reconstructed spectral images at higher resolution. We shall discuss the well foundedness of the above approach and describe our numerical approach. Finally, some experiments for 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 this image as the panchromatic image. A multispectral image may be represented by a function  $\vec{u}$  from  $\mathbb{R}^2$  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 to the colour channels 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, when the photonic flux is subjected to an spectral selective filter, be it in the visible range, the near infrared or the ultraviolet one. We shall 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 to compute the total spectral energy from the above channels, the corresponding panchromatic image will be given by

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

Let us mention that 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 in fact 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.

The purpose of this paper will be to describe a method to increase the resolution of satellite multispectral images when we know the corresponding panchromatic image at a higher resolution. We shall assume that the panchromatic image  $u$  has been sampled at higher resolution, and we have an image  $u(i, j)$  of size  $N \times N$  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  $s$  being explicitly included to stress the loss of resolution of the multispectral data. Our purpose will be to reconstruct the high resolution multispectral image  $\vec{u} = (u_1, \dots, u_m)$ , which will be an image of size  $N \times N$ , knowing the data  $u$  and  $\vec{u}^s$ . For that we shall take into account several constraints imposed by the data generation model. In particular, for each channel, the low resolution pixel is formed from the high resolution pixel by a low pass filtering followed by a subsampling. If we denote by  $k_r$  the impulse response of this filter corresponding to channel  $r \in \{1, \dots, m\}$ , we may write

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

Obviously, we have to respect also the relation (1). Finally, we shall use the geometric information contained in the panchromatic image. Indeed, we shall constraint the geometry of the spectral channels at higher resolution to follow the geometry of the panchromatic image. This constitutes the main feature of our approach and needs further explanation. First, we shall explain what do we understand by the geometric information contained in the panchromatic image, and more generally, in any scalar image, including any spectral channel. Then we shall explain the reasons which support the underlying assumption that *images taken on different spectral bands share common geometric information*. For that we shall review the main conclusions in [7],[12].

Before going into the details of our method, let us say that there exists a huge literature on the subject. The most performing 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 [16], or improvements based on multiresolution analysis of P+XS images [1]. For more information we refer to [1],[16], and references therein.

## 1.1 Mathematical morphology of scalar images

In this subsection, we shall consider scalar images, that is, images with a single channel, be it colour (or any other spectral channel) or grey level.

What is the geometric information content of an image ? This is the question we want to discuss briefly here. In this paragraph, we are simply summarizing some arguments contained explicitly or implicitly in the Mathematical Morphology theory [15], which were further developed in [6].

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 [17] to state as a basic principle that the gray level is not an observable.

Mathematical Morphology recognized contrast invariance as a basic invariance requirement and proposed that image analysis operations take into account this invariance principle [15, 11]. 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 in the image is contained in the level sets, independently of their actual levels. Thus, we are led to consider that the geometric information, the shape information, is contained in those level sets.

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 [2] 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 [3, 10]), 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* [6]. The topographic map is invariant under a wide class of local contrast changes ([6]), and, in particular, it is a useful tool for comparing images of the same object with different illuminations [12, 13]. We can conceive the topographic map as a tool giving a complete description of the geometry for 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 [2], we have that (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. This gives a unit vector field  $\theta$  defined at the points of  $\Omega$  which are traversed by a level line, and verifying  $\theta \cdot \nabla u = |\nabla u|$ , identity which, in a rigorous mathematical sense, can be understood as equality of two measures. This vector field has to be extended to all of  $\Omega$ . For that, we may define it to be zero at the flat regions of the graph of  $u$ , so that, essentially we have  $\theta(x) = \frac{\nabla u(x)}{|\nabla u(x)|}$  if  $\nabla u(x) \neq 0$  and  $\theta(x) = 0$  when  $\nabla u(x) = 0$ . A more precise description of  $\theta$  will be given in Section 2. This vector field will be the right analytic tool which we require to impose the constraint that the geometry of any function  $v$  is specified by the geometry of  $u$ , indeed, by  $\theta$  [5].

## 1.2 Geometry and color in natural images

What is 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 were 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 ? This problem was addressed in [7] and it was experimentally checked that the essential geometric contents of a color image is contained in the level lines of the corresponding intensity image.

Let us describe with some detail the experimental set up which permitted the authors of [7] to support the above conclusion. For that, we adopt the hypothesis that the essential geometric contents of a color image is contained in the level lines of its intensity. Then [7], an algorithm was designed 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 can be briefly described: *replace 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 [7] displayed several experiments confirming this hypothesis. A further confirmation of it was suggested by the experiments of imposing the color of an image to 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 [7].

## 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 a common geometric information. Indeed, this assumption

was experimentally studied in [12]. The experiments were done on images of four spectral channels corresponding to the blue, green, red and near infrared regions. Two channels were compared by means of 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 [12]. With both comparison procedures the conclusion was the same, the channels which are in the visible region share a large portion of its topographic map, this amount decreases, but it is still large for the near infrared channel. For the red and near infrared images, it was also showed that, after contrast inversion, there is still a portion of the topographic map which is common to the topographic map of the blue channel [12]. This was explained in terms of the vegetation, which become apparent in the red and near infrared channels by contrast inversion [12]. The methods used were robust under illumination changes and geometric transformations, in particular, under changes of scale [12].

## 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 the problem described above: increase the resolution of satellite multispectral images when we know the corresponding panchromatic image at higher resolution.

Let us explain the plan of the report. Section 2 is devoted to recall some basic facts about functions of bounded variation. Section 3 is devoted to describe the energy functional of our variational model. Section 4 is devoted to the description of the algorithm and the numerical experiments. We end up in Section 5 with some conclusions. Finally, in Section 6 we briefly comment on the mathematical justification of the problem.

## 2 Function spaces

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

Let us first recall the definition of  $BV$  functions and total variation. Let  $Q$  be an open set in  $\mathbb{R}^2$ . 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

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

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

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

In any case, if  $u \in BV(Q)$  we shall always write  $\|\nabla u\| = \int_Q |\nabla u|$ . The space  $BV(Q)$  is endowed with the norm

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

Let  $H^1$  denote the one-dimensional Hausdorff measure in  $\mathbb{R}^2$  (a measure of length). We say that 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. For sets of finite perimeter  $E$  one can define the essential boundary  $\partial^* E$ , which is rectifiable with finite  $H^1$  measure, and compute the normal to the level set at  $H^1$  almost all points of  $\partial^* E$ . Thus at almost all points of almost all level sets of  $u \in BV(Q)$  we may define a normal vector  $\theta(x)$ . 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,  $|\theta| \leq 1$  a.e.. We shall refer to the vector field  $\theta$  as the vector field of unit normals to the topographic map of  $u$ . For further information concerning functions of bounded variation we refer to [3, 10, 18].

### 3 The energy functional: continuous and discrete description

We shall 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, we 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  $(X_1, X_2, X_3, X_4)$  and call them the red, green, blue and near infrared channels, even if they could be different spectral channels. **Indeed in our experiments below they will represent the red, green, blue, and near infrared channels.** The extension to any set of spectral channels will be immediate. As above, we denote by  $u$  the intensity image corresponding to  $\vec{u}$ .

Assume that 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 by  $\vec{u}^S$  the known values of  $\vec{u}$  on  $S$ , in coordinates,  $\vec{u}^S = (X_1^S, X_2^S, X_3^S, X_4^S)$ . Recall that the low resolution pixel is formed from the high resolution pixel by a 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 the discussion of Section 1, we should impose the following two relations

$$u(x) = \alpha_1 X_1(x) + \alpha_2 X_2(x) + \alpha_3 X_3(x) + \alpha_4 X_4(x) \quad (7)$$

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

$$X_n^S(i, j) = k_n * X_n(i, j), \quad \forall (i, j) \in S, \quad n = 1, 2, 3, 4 \quad (8)$$

which correspond to (1), and (2), respectively. To give a sense to the relations in (8), we need to assume that it is possible to evaluate  $k_n * X_n$  at any point of  $S$ . For that, being  $k$  any of the kernels  $k_1, k_2, k_3, k_4$ , we shall assume that

(H)  $k$  is the kernel of a convolution operator mapping  $L^2(\Omega)$  into  $C(\overline{\Omega})$ .

Under assumption (H), for any point  $(i, j) \in \Omega$ , the map which to any  $f \in L^2(\Omega)$  associates the value

$$k * f(i, j) = \int_{\Omega} k((i, j) - (x, y))f(x, y) dx dy$$

is a linear functional in  $L^2(\Omega)$  and relations (8) have a 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 determine uniquely 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 we recalled in Section 2, there is a vector field  $\theta : \Omega \rightarrow \mathbb{R}^2$  with  $|\theta| \leq 1$  such that

$$\theta \cdot \nabla u = |\nabla u|$$

as measures in  $\Omega$ . Moreover

$$|\theta| = 1 \quad |\nabla u| \text{ a.e.}$$

The vector field  $\theta$  represents the vector field of unit normals to the level sets of  $u$ . In practice, at the discrete level,  $\theta$  can be defined by the relation  $\theta(x, y) = \frac{\nabla u(x, y)}{|\nabla u(x, y)|}$  when  $\nabla u(x, y) \neq 0$ , and  $\theta(x, y) = 0$  when  $\nabla u(x, y) = 0$ .

Given the vector field  $\theta$  of unit normals to the level sets of  $u$ , and assuming that  $X_1, X_2, X_3, X_4 \in BV(\Omega)$  we shall require that

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

The same relationship can be imposed in a slightly different way that will have important consequences mainly computational (since it will make the algorithm much faster) but also will permit to construct a functional which is invariant under contrast inversion, a suitable property when considering the case of false color images (see [12]). Since  $\theta$  has the direction of the normal to the level lines of  $u$ , the counterclockwise rotation of angle  $\pi/2$ , denoted by  $\theta^\perp$ , represents the tangent vector to the level lines of  $u$ . In this case, if the spectral channels share the geometry of the panchromatic image, we have

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

Obviously, the relations (9) (also relations (10)) cannot be exactly satisfied and we have to impose them in a variational framework (together with the other constraints discussed at the introduction) by minimizing the sum of integrals

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

In a similar way, the relations (10) can be imposed in a variational framework by minimizing the sum of integrals

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

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

Let us write functional (11) in a convenient form for computational purposes. For that, we integrate by parts the second term of each integral, and 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

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

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

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

We may impose (8) by minimizing the sums

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

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

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

for each  $n = 1, 2, 3, 4$ , where  $X_n^S(x,y)$  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 term (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 a mathematical justification of the algorithm.

**The energy functional.** Thus, we propose to compute the high resolution multispectral images  $X_1, X_2, X_3, X_4$  by minimizing the energy functional:

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

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), and  $p = 1$ , or 2. For the purposes of comparison, let us also write a variant of (19) which is based on (13)

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

subject to the same constraints as (19). Observe that the first integral of (20) could also be written in the form (11). Experiments based on this functional have been reported in [4]. Finally, let us note that the first integral in (20) was also used in [5] in the context of filling-in by joint interpolation of vector fields and gray levels, in order to constrain the vector field  $\theta$  and the image  $u$  to be related by  $\theta \cdot \nabla u = |\nabla u|$ , but, in this case, both  $\theta$  and  $u$  were unknown.

In both functionals we take  $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \lambda, \mu > 0$  (in practice, all these parameters are taken equal to 1), and  $p = 1$ , or 2. In case we consider only the channels  $X_1 = \text{red}$ ,  $X_2 = \text{green}$ ,  $X_3 = \text{blue}$ , both functionals can be effectively used as we have shown [4]. In case that we also add the channel  $X_4 = \text{near infrared}$ , it seems slightly better to use functional (19), or a variant of it which uses (10) for the near infrared channel and (9) for the red, green, and blue channels. Notice 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 functional (19) is that the case  $p = 2$  gives Euler-Lagrange equations which are linear in the  $X_n$  variables and a steepest descent converges to the minimum in a much faster way than (20). Indeed, for an image of size  $800 \times 800$  the execution time is of 63 seconds in a on a Pentium 1.8 GHz.

### 3.2 The discrete formulation

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

$$\nabla^{+,+} f = (\nabla_x^+ f, \nabla_y^+ f), \quad \nabla^{+,-} f = (\nabla_x^+ f, \nabla_y^- f), \quad \nabla^{-,+} f = (\nabla_x^- f, \nabla_y^+ f), \quad \nabla^{-,-} f = (\nabla_x^- f, \nabla_y^- f)$$

where

$$\nabla_x^+ f(i, j) = f(i+1, j) - f(i, j), \quad \nabla_x^- f(i, j) = f(i, j) - f(i-1, j),$$

$$\nabla_y^+ f(i, j) = f(i, j+1) - f(i, j), \quad \nabla_y^- f(i, j) = f(i, j) - f(i, j-1).$$

Note that the dual operators to  $\nabla^{+,+}$ ,  $\nabla^{+,-}$ ,  $\nabla^{-,+}$ ,  $\nabla^{-,-}$  are, respectively, the operators  $\operatorname{div}^{-,-}$ ,  $\operatorname{div}^{-,+}$ ,  $\operatorname{div}^{+,-}$ ,  $\operatorname{div}^{+,+}$ . 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  $\alpha, \beta = +, -$ .

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

$$\begin{aligned}
& \sum_{n=1}^4 \frac{\gamma_n}{4} \sum_{\alpha, \beta = +, -} \sum_{i, j=0}^{N-1} |\theta_{\alpha, \beta}(i, j)^\perp \cdot \nabla^{\alpha, \beta} X_n(i, j)|^2 + \\
& \lambda \sum_{i, j=0}^{N-1} \left( \sum_{n=1}^4 \alpha_n X_n(i, j) - u(i, j) \right)^2 + \\
& \mu \sum_{n=1}^4 \sum_{(i, j) \in S} \left( (k_n * X_n(i, j) - X_n^S(i, j))^2 \right),
\end{aligned} \tag{21}$$

subject to  $0 \leq X_n \leq M_n$ ,  $n = 1, 2, 3, 4$ .

Observe that we have used simultaneously 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

For simplicity, we shall only describe the algorithm used to minimize (21). Applying the gradient descent method, to minimize (21) we iteratively actualize the solution using the equations

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

where  $n = 1, 2, 3, 4$ ,  $p \geq 0$ . 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 to write in detail the terms involving  $k_n$ . In practice, it is very important to guarantee that the energy, call it  $E_\epsilon(X_1^p, X_2^p, X_3^p, X_4^p)$ , decreases along the evolution, i.e., that

$$E_\epsilon(X_1^{p+1}, X_2^{p+1}, X_3^{p+1}, X_4^{p+1}) \leq E_\epsilon(X_1^p, X_2^p, X_3^p, X_4^p). \tag{23}$$

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 to choose  $\Delta t$  so that the energy has the largest decreasing in the direction of the energy gradient. If  $\nabla E_\epsilon(X_1^p, X_2^p, X_3^p, X_4^p)$  denotes the energy gradient which is given in the right-hand side of (22), then  $\Delta t$  can be chosen as a solution of

$$\min_{s>0} E_\epsilon((X_1^p, X_2^p, X_3^p, X_4^p) - s \nabla E_\epsilon(X_1^p, X_2^p, X_3^p, X_4^p)). \tag{24}$$

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

### 4.1 Description of the data

To test our method we shall dispose of some reference multispectral satellite images, which have been furnished to us by the CNES for this purpose. The data consists of a panchromatic image  $u(i, j)$ ,  $i, j \in \{0, \dots, N-1\}$ , at

resolution  $0.7 \text{ m/pixel}$ , and a set of multispectral images

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

at resolution  $2.8 \text{ m/pixel}$ , which represent the red, green, blue, and near infrared spectral channels. To test our experiments, we also dispose of a reference image

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

at resolution  $0.7 \text{ m/pixel}$ .

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

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

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

In the experiments displayed in this paper, the panchromatic constraint (7) is

$$u = 0.5X_1 + 0.5X_2, \tag{25}$$

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  $41 \times 41$ .

The range of values of the red, green, blue, and near infrared channels data is the interval  $[0, 512]$ .

## 4.2 Numerical experiments

We display some experiments on a set of multispectral satellite images made of four channels. In the experiments below we have tested an increase of resolution by a factor of 4 and we displayed our results in the case of true color (where we display the red, green, and blue channels) and in case of false color (where we display the red, green, and near infrared channels).

We iteratively minimize (21) using the gradient descent equations given in (22). We take as initialization of (22) an image obtained by simple replication of  $\vec{u}^{2.8}$  by a factor 4. Other initializations like a DCT zoom (of factor 4), or a more sophisticated initialization obtained using the approach in [8] can also be taken, and could be useful to speed up the algorithm. In our experiments we have checked that the results obtained are similar.

$u$  and  $\vec{u}^{2.8}$  constitute the basic data of the problem. We stress the fact 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 shall first display the true color images (involving the RGB channels) and then the false color images (involving the R,G,IR 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  $(R_{ref}, G_{ref}, B_{ref}, IR_{ref})$  both visually and by displaying some error measures. We shall compute the errors on the whole image and on certain representative regions (see Figure 10). Let  $X$  denote the region where the error is computed (the whole image or a region of it). We shall use the maximum, the  $\ell^1$  and  $\ell^2$  error measures, i.e.,



Figure 1: the true color reference image  $(R_{ref}, G_{ref}, B_{ref})$ .

(i) The maximum error:

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

(ii) The  $\ell^1$  norm

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

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

(iii) The  $\ell^2$  norm

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

with similar expressions for the other channels. We also display the histogram of errors, and we make explicit some percentiles.

#### 4.2.1 Display of true color results

Figure 1 displays the red, green, and blue channels of reference image  $\vec{u}^{ref}$ , i.e.,  $(R_{ref}, G_{ref}, B_{ref})$ .

Figure 2 displays the panchromatic image  $u$ , and Figure 3 the RGB spectral channels  $(R_{2.8}, G_{2.8}, B_{2.8})$  of the data at lower resolution  $\vec{u}^{2.8}$ . Both images have been furnished to us by the CNES.

Figure 5 displays the  $(R, G, B)$  channels of the reconstructed image at the resolution of the panchromatic image obtained using functional (21). The  $k_n * X_n$  convolution has been computed in the Fourier domain. The result has to be compared with the reference image, Figure 1. The RGB channels of the initialization of (22) has been displayed in Figure 4.

Table 11 displays the errors between the reference image and the result, in particular the errors in the RGB channels. Below the table, we display the corresponding histogram of errors.



Figure 2: The panchromatic image at resolution 0.7 m/pixel, denoted in the text by  $u$ .



Figure 3:  $(R_{2.8}, G_{2.8}, B_{2.8})$  spectral channels at resolution 2.8 m/pixel, i.e., the RGB coordinates of  $\vec{u}^{2.8}$ .

#### 4.2.2 Display of false color results

False color images are composed by red, green, and near infrared channels. Figure 6 displays the  $(R_{ref}, G_{ref}, IR_{ref})$  channels of the reference image  $\vec{u}^{ref}$ . Figure 7 displays the spectral channels  $(R_{2.8}, G_{2.8}, IR_{2.8})$  at resolution 2.8 m/pixel.

Figure 9 displays the  $(R, G, IR)$  channels of the reconstructed image at the resolution of the panchromatic image obtained using functional (21). The  $(R, G, IR)$  channels of the initialization are displayed in Figure 8.

Table 11 displays the errors between the reference image and the result, in particular the errors in the IR channel. Below the table, we display the corresponding histogram of errors.

## 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 the low resolution pixel is formed from the high resolution pixel by a low pass filter followed by subsampling (2). But 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*



Figure 4: RGB channels of the initialization (obtained by replication by a factor 4) of the algorithm using functional (19) with  $p = 2$ .



Figure 5: the RGB channels of the reconstructed image (at the resolution of the panchromatic image) obtained using functional (19) with  $p = 2$ .

*panchromatic image*. We have constructed two slightly different energy functionals (20) and (19) which incorporate the above three basic postulates. We have to 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 them and, for simplicity, we displayed some experiments relative to functional (19) showing its capabilities in true color and false color.



Figure 6: the  $(R_{ref}, G_{ref}, IR_{ref})$  channels of the reference image  $\bar{u}^{ref}$ , which display the false color reference image.



Figure 7:  $(R_{2.8}, G_{2.8}, IR_{2.8})$  spectral channels at resolution 2.8 m/pixel, i.e., the  $R, G, IR$  channels of  $\bar{u}^{2.8}$ .

The results with both functionals are comparable.

Finally, let us mention that the Euler-Lagrange equations of functional (19) with  $p = 2$  are linear in the variables and the gradient descent exhibits a fast convergence. The number of operations per pixel is 2700. The time spent reconstructing an image of size 800x800 pixels on a Pentium 1.8 GHz is 63 seconds. In the cases where functional (20) produces better results we could compute first the result with functional (19) with  $p = 2$  and then do some iterations with functional (20) to improve the result.

**Acknowledgement.** This work was supported by the CNES.

## 6 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  $div \theta^\perp \in L^2(\Omega)$ . Let  $W(\Omega, \theta)$  be the completion of  $W^{1,2}(\Omega)$  with respect to the norm

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

We have the following result.



Figure 8:  $R, G, IR$  channels of the initialization (obtained by replication by a factor 4) of the algorithm using functional (19) with  $p = 2$ .



Figure 9: the  $R, G, IR$  channels of the reconstructed image (at the resolution of the panchromatic image) obtained using functional (19) with  $p = 2$ .

**Theorem 1** *If  $\text{div } \theta^\perp \in L^2(\Omega)$ , then the functional (19) admits a minimum in  $W(\Omega, \theta)^3$ . Similarly, if  $\text{div } \theta \in L^2(\Omega)$ , functional (20) admits a minimum in  $BV(\Omega)^3$ .*

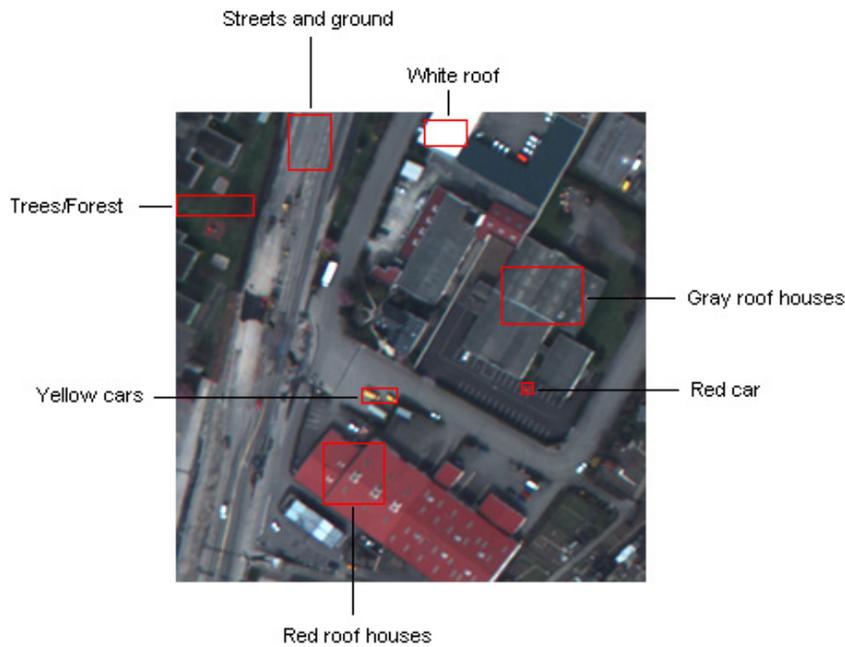


Figure 10: the regions where the statistics of Table 11 are computed.

## References

- [1] 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. on Geoscience and Remote Sensing*, vol. 40, n. 10, pp. 2300–2312, 2002.
- [2] L. Ambrosio, V. Caselles, S. Masnou and J.M. Morel, The Connected Components of Sets of Finite Perimeter, *European Journal of Math.*, 3, pp. 39-92, 2001.
- [3] L. Ambrosio, N. Fusco and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Mathematical Monographs, 2000.
- [4] C. Ballester, V. Caselles, B. Rougé, and J. Verdera, *Une méthode géométrique de fusion des images P+XS*. Rapport CNES, 2003.
- [5] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, *Filling-In by Joint Interpolation of Vector Fields and Gray Levels*, *IEEE Transactions on Image Processing*, Vol.10, n. 8, pp.1200-1211, 2001.
- [6] V. Caselles, B. Coll and J.M. Morel, *Topographic Maps and Local Contrast Changes in Natural Images*, *Int. J. Comp. Vision* 33(1), pp. 5-27, 1999.
- [7] V. Caselles, B. Coll and J.M. Morel, *Geometry and Color in Natural Images*, *J. Math. Imaging and Vision*, 16, pp. 89-107, 2002
- [8] V. Caselles, G. Sapiro and D.H. Chung, *Vector Median Filters, Vector Morphology, and Coupled PDE's: Theoretical Connections*, *Journal of Mathematical Imaging and Vision*, 12, pp. 109-120 (2000).
- [9] S. Durand, F. Malgouyres and B. Rougé, *Image Deblurring, Spectrum Interpolation and Application to Satellite Imaging*, *Mathematical Modelling and Numerical Analysis*, 1999.
- [10] L.C. Evans and R.F. Gariepy, *Measure Theory and Fine Properties of Functions*, *Studies in Advanced Math.*, CRC Press, 1992.

	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

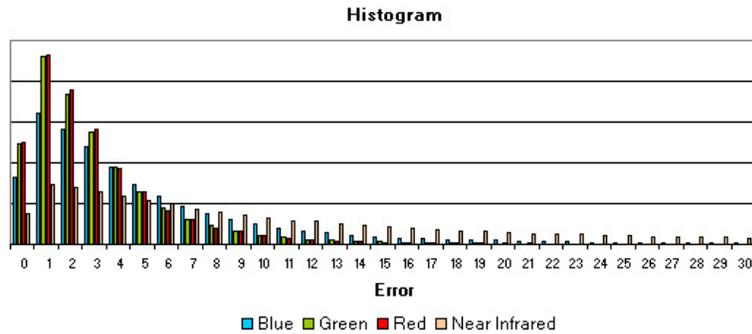


Figure 11: Errors corresponding to the experiment above.

- [11] F. Guichard and J.M. Morel, *Partial differential equations and image iterative filtering*, Forthcoming book.
- [12] J.L. Lisani and J.M. Morel, *Comparaison morphologique d'images*, Rapport CNES, 2000.
- [13] P. Monasse, *Contrast invariant image registration*, Proc. of the International Conf. on Acoustics, Speech and Signal Processing, Phoenix, Arizona, **6** (1999), 3221–3224.
- [14] L.I. Rudin, S. Osher and E. Fatemi, *Nonlinear Total Variation Based Noise Removal Algorithms*. *Physica D*, **60**, 259–269.
- [15] J. Serra, *Image analysis and mathematical morphology*. Academic Press, 1982.

- [16] L. Wald, [View publication stats](#) T. Ranchin, and M. Mangolini, Fusion of Satellite Images of Different Spatial Resolutions: Assessing the Quality of Resulting Images, *Photogramm. Eng. Remote. Sensing*, vol. 63(6), pp. 691–699, 1997.
- [17] M. Wertheimer, Untersuchungen zur Lehre der Gestalt, II, *Psychologische Forschung*, 4, 1923, pp. 301–350.
- [18] W.P. Ziemer, *Weakly Differentiable Functions*, GTM 120, Springer Verlag, 1989.