

A Variational Approach for Sharpening High Dimensional Images*

Michael Möller[†], Todd Wittman[‡], Andrea L. Bertozzi[‡], and Martin Burger[†]

Abstract. Earth-observing satellites usually not only take ordinary red-green-blue images but also provide several images including the near-infrared and infrared spectrum. These images are called multispectral, for about four to seven different bands, or hyperspectral, for higher dimensional images of up to 210 bands. The drawback of the additional spectral information is that each spectral band has rather low spatial resolution. In this paper we propose a new variational method for sharpening high dimensional spectral images with the help of a high resolution gray-scale image while preserving the spectral characteristics used for classification and identification tasks. We describe the application of split Bregman minimization to our energy, prove convergence speed, and compare the split Bregman method to a descent method based on the ideas of alternating directions minimization. Finally, we show results on Quickbird multispectral as well as on AVIRIS hyperspectral data.

Key words. image fusion, multispectral, hyperspectral, pan-sharpening, variational

AMS subject classifications. 94A08, 65K10, 86-08, 68U10

DOI. 10.1137/100810356

1. Introduction. Instead of just taking red-green-blue images, many satellite imaging systems, such as the Quickbird and Landsat-7 satellites, produce so-called multispectral images, including the near-infrared spectrum, and consist of four to seven bands. These additional bands can be used for various identification and classification tasks. Because images in a precise spectral range can be taken at only rather low spatial resolution, many satellites also include a so-called panchromatic image, which is a gray-scale image that spans a wide range of frequencies but comes at high spatial resolution. Pan-sharpening is the process of fusing the low resolution multispectral image with the high resolution panchromatic image to obtain a high resolution multispectral image.

The goal of pan-sharpening is to combine the high spatial resolution of the panchromatic image with the precise spectral information of the multispectral image. The resulting image should have high visual quality to aid in detection and classification tasks. However, the pan-sharpened image should also contain the same spectral (color) information as the original multispectral data for precise identification of targets.

*Received by the editors October 1, 2010; accepted for publication (in revised form) November 11, 2011; published electronically January 24, 2012. This work was supported by the US Department of Defense, ONR grant N000140810363, NSF grant ACI-0321917, and NSF grant DMS-0601395. Preliminary computational results on hyperspectral data were previously presented at the SPIE Conference on Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV [35]. This paper is based on the results from the masters thesis [34].

<http://www.siam.org/journals/siims/5-1/81035.html>

[†]Institut für Numerische und Angewandte Mathematik, Westfälische Wilhelms-Universität Münster, Einsteinstr. 62, D 48149 Münster, Germany (m.moeller@gmx.net, martin.burger@uni-muenster.de). The last author's work was supported by DFG project "Regularization with Singular Energies."

[‡]Department of Mathematics, University of California, Los Angeles, CA 90099 (wittman@math.ucla.edu, bertozzi@math.ucla.edu).

Several methods have been proposed for pan-sharpening multispectral imagery. Many techniques express the panchromatic image as a linear combination of the multispectral bands, including the intensity-hue-saturation (IHS) [17, 16, 45] and Brovey methods [22]. Other methods such as principal component analysis (PCA) project the images into a different space [43]. Several authors have proposed using the wavelet transform or other types of multiresolution analysis (MRA) to extract geometric edge information from the panchromatic image. These details are injected into the low resolution image, sometimes based on certain decision models [51, 30, 33, 1]. Improvements, hybrid schemes, and other methods focusing on spectral consistency have been proposed [2, 42, 38, 24, 26]. Recently, Ballester et al. introduced a variational method called P+XS image fusion that explicitly forces the edges of the pan-sharpened image to line up with those in the panchromatic image [5].

Several survey papers have compared the performance of pan-sharpening methods [47, 21, 22, 4]. Generally, the methods producing the highest spatial quality tend to distort the spectral quality of the image. Very interesting observations have been made by Thomas et al. in [44]. One of the main points in their paper is that imposing a relationship between the remotely sensed images comes with the risk of drastically weakening the fusion performance.

Almost all methods impose a certain spectral relation between the panchromatic image and the spectral bands like the common assumption that the panchromatic image can be modeled as a linear combination of the spectral bands. As described in [44], this assumption can lead to heavy spectral distortion, since it does not reflect the physics correctly. Furthermore, these methods do not extend to higher dimensional images. For bands whose frequency is not covered by the panchromatic image, their main assumption is no longer valid.

Our goal is to combine ideas from wavelet fusion and P+XS fusion with a new spectral consistency term into a single variational framework that does not impose any direct spectral relation between the multispectral and the panchromatic images. We should mention that our approach does not respect all possible image input characteristics described in [44], but no assumption is made on the link between the intensity values of the multispectral and the panchromatic images. In other words, the high resolution gray-scale image does not have to be panchromatic but could be any high resolution image, an image we call the master image, showing the same scene as the multispectral image. Therefore, our method naturally extends to higher dimensions, i.e., hyperspectral images. Despite the missing spectral relation, we still want to preserve the spectral information from the low resolution spectral image in the sharpening process. We will present two slightly different energies—one which explicitly includes matching of wavelet coefficients in the wavelet domain and an alternate energy which can be minimized in the spatial domain only and therefore has big computational advantages. We refer to our models as variational wavelet pan-sharpening (VWP) and the alternate energy (AVWP).

In section 2, we present the energy models that describe the desired qualities of the sharpened image. We discuss numerical methods for minimizing the energy in section 3. In section 4, we present results on multispectral Quickbird data as well as hyperspectral data which we sharpen with Google Maps [27] screen shots as master images. We also show how to preserve the spectral information from the original data. Finally, we conclude in section 5 by suggesting extensions of VWP and areas for future research.

A more detailed description of our image fusion method can be found in the Master's thesis [34]. Preliminary results on hyperspectral data were previously presented at the SPIE Conference on Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV [35].

2. Energy functional. The general idea of variational image processing methods is to develop an energy functional depending on an image, where a low value of the energy functional corresponds to a good quality image.

Many existing pan-sharpening methods like IHS, Brovey, and P+XS image fusion assume that the panchromatic image is a linear combination of the different bands. Looking at the spectral response of the sensors of the Quickbird satellite system, this assumption does not seem to be true in general [38, Figure 1]. A false linear combination assumption can lead to spectral distortion and therefore damage the spectral information of the original multispectral image [44]. Furthermore, imposing a relation between the high resolution and the multispectral image as well as modeling the image formation process makes the method sensor dependent and restricts it to the bands whose spectral response is covered by the master image.

Therefore, we propose a method that does not depend on this assumption. The VWP energy model consists of four parts and is specifically designed to sharpen images with an arbitrary number of bands while preserving spectral quality. In the following we describe each term of our energy functional separately. We refer to the master image as $M : \Omega \rightarrow \mathbb{R}$, where $\Omega \subset \mathbb{R}^2$ is the image domain. H_i are the original low resolution and u_i the desired high resolution spectral bands.

2.1. Geometry enforcing term. The first variational pan-sharpening method was proposed by Ballester et al. in 2006 [5]. Their idea for introducing the geometry of the panchromatic image was to align all level lines of the high resolution image with each multispectral band. The main assumption here is that the geometric information of an image is contained in its level sets, independent of their actual level.

The level sets of an image can be represented by the vector field θ consisting of all unit normal vectors of those level sets. The definition of θ is somehow problematic: On smooth parts of the image the gradient is zero, and at edges a classical or a weak derivative does not exist. For practical purposes we use an additional regularization and define $\theta_\epsilon(x) = \frac{\nabla M(x)}{|\nabla M(x)|_\epsilon}$, where $|\nabla M|_\epsilon = \sqrt{(D_x M)^2 + (D_y M)^2 + \epsilon^2}$ is an approximation of the unit normal vector field.

For the true unit normal vector field θ we have $\theta \cdot \nabla M = |\nabla M|$. To ensure that each spectral band has the same level sets as the panchromatic image, Ballester et al. suggested aligning the normal vectors of the level sets. Therefore, every band of the restored image should satisfy $|\nabla u_n| - \theta \cdot \nabla u_n = 0$. Integrating over the left-hand side and applying integration by parts leads to $\int_{\Omega} (|\nabla u_n| + \operatorname{div}(\theta) u_n)$. The total energy Ballester et al. suggested to minimize in [5] is

$$\begin{aligned} J_{P+XS}(u) = & \sum_{n=1}^N \int_{\Omega} (|\nabla u_n| + \operatorname{div}(\theta) u_n) dx \\ & + \lambda \int_{\Omega} \left(\sum_{n=1}^N \alpha_n u_n - M \right)^2 dx \end{aligned}$$

$$(2.1) \quad +\mu \sum_{n=1}^N \int_{\Omega} \pi_S((k * u_n) - H_n)^2 dx,$$

where λ and μ are parameters, k is a convolution kernel modeling the blur of the low resolution data, and π_S is a Dirac comb indicating which multispectral pixel values are given by the low resolution data. Increasing the resolution by a factor of four, the discrete Dirac comb would have ones every fourth pixel and zeros elsewhere. This energy was proposed for sharpening multispectral images with M being a high resolution panchromatic image. The second term reflects the assumption that M is a linear combination of the spectral bands u_n with mixing coefficients α_n .

The idea for introducing the geometry of a higher resolution image can easily be adapted for the sharpening of hyperspectral images with an arbitrary master image at high resolution. θ depends only on the level sets of the image, and therefore no assumption on the relation between the hyperspectral image and the master image needs to be made. In particular, the method is independent of contrast and illumination changes in the master image. The inversion of the contrast, however, does change the geometry introducing term, which is why Ballester et al. furthermore proposed a term of the form

$$(2.2) \quad \int_{\Omega} |\nabla u_n \cdot \theta^\perp|$$

to obtain invariance under contrast inversion.

We start the construction of our energy similarly to the P+XS model. As discussed above, introducing the geometric information from the master image means minimizing the quantity $|\nabla u_n| - \theta \cdot \nabla u_n$ over the whole image domain Ω . Integration by parts leads to

$$(2.3) \quad \hat{E}_g = \gamma \sum_{n=1}^N \left[\int_{\Omega} |\nabla u_n| dx + \int_{\Omega} \operatorname{div}(\theta) \cdot u_n dx \right].$$

Although (2.3) is not invariant under contrast inversion, we obtained better results in our numerical experiments with this term in comparison to (2.2), which we think is due to regularizing the L^1 norm of the gradient of u_n rather than a quadratic deviation of θ . Furthermore, the term (2.3) has a nice link to the recent work in image denoising and Bregman iteration: The idea of aligning the gradient vectors of an image with a smooth unit normal vector field similar to (2.3) was also proposed by Lysaker, Osher, and Tai [31] for a new image denoising algorithm and led to iterative regularization using the Bregman distance [37]. Therefore, it is not surprising that the geometry introducing term (2.3) also has a link to Bregman distances. In the next subsection we will briefly recall the concept of Bregman distances and show the connection to the contour alignment term (2.3).

2.1.1. Bregman distance and its link to the alignment of isocontours. The (generalized) Bregman distance associated with a convex energy functional J is defined as

$$(2.4) \quad D_J(u, v) = \{J(u) - J(v) - \langle p, u - v \rangle \mid p \in \partial J(u)\},$$

where $\partial J(u)$ denotes the subdifferential of J at u . While for a Fréchet differentiable functional the Bregman distance is unique, it can be multivalued for nondifferentiable functionals like the total variation (TV). Each element $D_J^p(u, v)$ represents a distance in the sense that $D_J^p(u, v) \geq 0$. For a strictly convex J we even get $D_J^p(u, v) > 0$ for $u \neq v$.

Bregman distances were originally introduced in [6] as a functional analytical concept to find extrema of convex functionals. Recently, many papers have made use of Bregman distances to develop new image processing and optimization algorithms and prove convergence rates (e.g., [37, 7, 25, 9]).

It turns out that in the case of the geometry enforcing term from P+XS, (2.3), we are also minimizing a Bregman distance: For a smooth master image with nonzero gradient, $-\operatorname{div}(\theta) = -\operatorname{div}(\frac{\nabla M}{|\nabla M|})$ is the subgradient of the TV seminorm at M . Hence, we know that $|M|_{TV} = \langle -\operatorname{div}(\theta), M \rangle$. Adding zero in the form of $-|M|_{TV} + \langle \operatorname{div}(\theta), -M \rangle$ to term (2.3) then gives us

$$(2.5) \quad \begin{aligned} \hat{E}_g &= \gamma \sum_{n=1}^N \left(|u_n|_{TV} - |M|_{TV} + \langle \operatorname{div}(\theta), u_n - M \rangle \right), \\ &= \gamma \sum_{n=1}^N D_{|.|_{TV}}^{-\operatorname{div}(\theta)}(u_n, M), \end{aligned}$$

which shows that aligning the isocontours of each multispectral band with the master image is nothing but minimizing their Bregman distance with respect to the TV seminorm. We penalize derivations from the master image in the Bregman distance. It is shown in [40] that Bregman distance regularization with respect to TV is linked to texture enhancement. Furthermore, the Bregman distance is contrast invariant. This indicates that this type of regularization is very well suited for our problem because we are “close” to the master image independent of the master image’s actual values, which means that the edges and jumps in the image must be in the right places.

2.2. Contrast increasing contour alignment. Despite the nice link to Bregman distances, we shall also consider a different weighting of the two terms in the energy functional and write

$$(2.6) \quad E_g = \sum_{n=1}^N \left[\gamma \int_{\Omega} |\nabla u_n| dx + \eta \int_{\Omega} \operatorname{div}(\theta) \cdot u_n dx \right].$$

The introduction of the additional parameter η might not be immediately clear since it slightly contradicts the derivation of this term, i.e., that ideally we have $|\nabla u_n| - \theta \cdot \nabla u_n = 0$. We found during our numerical experiments that choosing a slightly higher parameter $\eta > \gamma$ gives better spatial quality. In the next paragraph we will look at the meaning of using different parameters in a mathematical sense and further motivate our decision by some simple one dimensional examples.

The optimality condition for minimizing the functional (2.6) is

$$(2.7) \quad \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) = \frac{\eta}{\gamma} \operatorname{div}(\theta).$$

First, note that the divergence of the normals (a quantity of the form $\operatorname{div}\frac{\nabla a}{|\nabla a|}$) has a geometric meaning, i.e., the mean curvature of the level lines. The optimality condition yields that for $\eta = \gamma$ we enforce the curvature of our desired image u to be the same as in our high resolution master image M . If we now choose η to be greater than γ , we enforce the curvature of the desired image u to be even higher. The desired curvature is the one of the panchromatic image multiplied by $\frac{\eta}{\gamma} > 1$. To see the effect of a higher curvature we implemented an evolution of the form

$$(2.8) \quad u^{k+1} = u^k + \tau \left(\operatorname{div}\left(\frac{\nabla u^k}{|\nabla u^k|}\right) - \frac{\eta}{\gamma} \operatorname{div}(\theta) \right)$$

for some example signals in one dimension where we took the data as our initialization u_0 and ran (2.8) for a limited number of iterations. The results are shown in Figure 1. One can see that all results (green curves) have a higher contrast than the original signals. The peaks and parts of high curvatures have larger values after the evolution. For the second example, the step function, the result can be proved analytically even for higher dimensions.

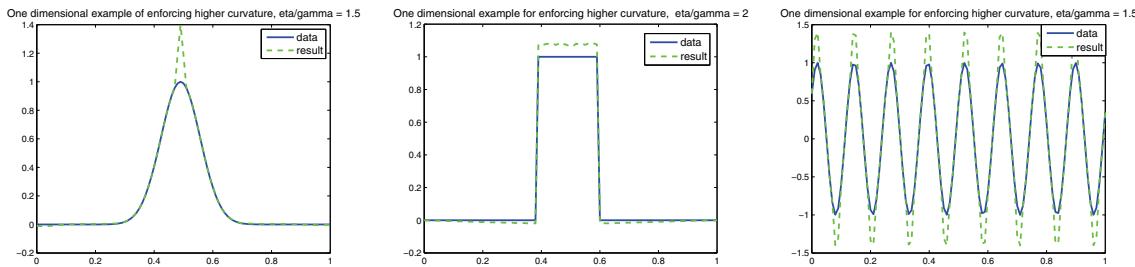


Figure 1. Examples of enforcing higher curvature. Left: Gaussian signal. Middle: step function. Right: sinusoidal signal.

In image denoising regularization by TV is known to decrease the contrast of images, while the weighting increases the contrast. Maybe the two effects cancel out such that, in sum, a better, more realistic image is obtained.

2.3. Fidelity terms. To preserve as much spectral information as possible we would like our fused spectral image to be very close to the low resolution image on those parts of the image where there are no edges or texture. On the edges we would like to increase the contrast and enhance the spatial information. Therefore, we decided to include a matching term in the wavelet domain.

A second level undecimated wavelet decomposition of the master image and each spectral band is calculated. A well-known concept in wavelet pan-sharpening is to use the detail coefficients of the high resolution image (most likely containing the geometric information) and the approximation coefficients of each multispectral band (containing the color or spectral information) to obtain a sharpened image via the reconstruction from these coefficients [51]. This concept is illustrated in Figure 2 for a stationary wavelet transformation.

Despite the higher computational costs as well as memory requirements, stationary wavelets have proved to be superior to the orthogonal wavelet transform because they are translational

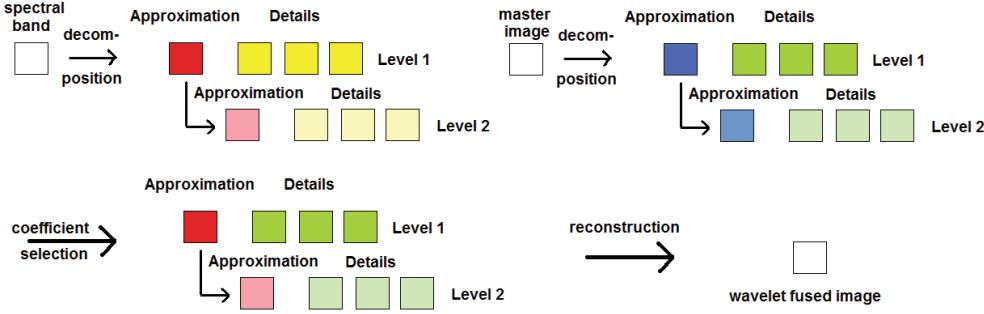


Figure 2. Concept of the wavelet pan-sharpening.

invariant and therefore much more robust toward slightly inaccurate registration and aliasing effects [15, 38, 28, 1].

In our energy we would like to be able to weight the importance of the wavelet coefficient matching differently for different levels of decomposition and types of coefficients. Depending on the desired image features, the matching of the detail coefficients from the high resolution image could be more or less important than the approximation coefficient matching to the low resolution image.

To formalize the wavelet matching in a mathematical context for our energy functional, let us briefly recall the wavelet notation. For a one dimensional wavelet transform, let ϕ be a scaling function and ψ the corresponding wavelet generating a wavelet orthonormal basis of $L^2(\mathbb{R})$. We define the wavelets $\psi^1(x) = \psi(x_1)\phi(x_2)$, $\psi^2(x) = \phi(x_1)\psi(x_2)$, $\psi^3(x) = \psi(x_1)\psi(x_2)$ and denote for $1 \leq k \leq 3$, $j \in \mathbb{Z}$, and $n = (n_1, n_2) \in \mathbb{Z}^2$

$$(2.9) \quad \psi_{j,n}^k(x) = \frac{1}{2^j} \psi^k \left(\frac{x_1 - 2^j n_1}{2^j}, \frac{x_2 - 2^j n_2}{2^j} \right).$$

Further we define a two dimensional scaling function by

$$(2.10) \quad \phi_{j,n}^2(x) = \frac{1}{2^j} \phi \left(\frac{x_1 - 2^j n_1}{2^j} \right) \phi \left(\frac{x_2 - 2^j n_2}{2^j} \right).$$

Then the approximation coefficients of a two dimensional function are given by the scalar product with ϕ^2 , and the three detail coefficients (which can be seen as horizontal, vertical, and diagonal details) are given by the scalar product with ψ^k , $k \in \{1, 2, 3\}$. We define the approximation matching coefficients for band i as

$$(2.11) \quad a_j^i[n] = \langle \uparrow H_i, \phi_{j,n}^2 \rangle,$$

where $\uparrow H_i$ denotes upsampling of the low resolution spectral band. In our experiments we used bilinear interpolation. The matching detail coefficients are taken from the scalar product with the panchromatic image and are equal for all different bands,

$$(2.12) \quad d_{\{k,j\}}[n] = \langle P, \psi_{j,n}^k \rangle \quad \text{for } 1 \leq k \leq 3.$$

If we denote the desired approximation coefficients for band i by $\alpha_j^i[n]$ and the desired detail coefficients by $\beta_{\{k,j\}}^i[n]$, then we add the following term to our energy functional:

$$(2.13) \quad E_w = \sum_n c_0 (a_L^i[n] - \alpha_L^i[n])^2 \phi_{j,n}^2(x) \\ + \sum_n \sum_{j=1}^L \sum_{k=1}^3 c_j (d_{\{k,j\}}[n] - \beta_{\{k,j\}}^i[n])^2 \psi_{j,n}^k(x),$$

where c_0 is the parameter for the approximation coefficient matching, c_j , $1 \leq j \leq L$, are the parameters for the different levels of detail coefficient matching, and L is the level of decomposition. In our experiments we used $L = 2$. Notice that we assumed that our continuous representations of the images are elements of $V_0^2 = \text{span}(\phi_{0,n}^2)$ since this holds for the discrete formulation anyway. In our variational context we choose the parameters according to the type of image we would like to produce. For high spatial quality we want to introduce the edge information of the panchromatic image and therefore increase c_1 and c_2 . Conversely, we would choose larger values for c_0 for better spectral quality.

A second matching term we introduce is based on the assumption that the best spectral information can be obtained from the low resolution image. Therefore, homogeneous parts without edges or texture should be contained as they are. However, the set of edges and texture is easy to get from the master image. Unlike segmentation methods (such as Mumford–Shah), where the edge set is an unknown, we can simply calculate it in the master image. To detect edges and texture we apply the exponential edge detector $\exp(-\frac{d}{|\nabla M|^2})$ to our master image M (with an appropriate constant d , e.g., $d = 0.004$). We add

$$(2.14) \quad E_c = \nu \sum_{i=1}^N \int_{\Omega \setminus \Gamma} (u_i - \uparrow H_i)^2 dx$$

to our energy functional.

The combination of terms in the spatial domain and in the wavelet domain has been applied in various image processing applications (e.g., [13, 19, 32]). Except for the advantage that one can choose different weights in the wavelet matching to obtain a certain type of image, this kind of combination has a drawback. Each minimization method will have to alternate between the spatial domain and the wavelet domain to include the different types of terms our energy functional contains. Using stationary wavelets this procedure can become extremely time consuming as well as memory consuming, particularly for very high dimensional images with up to 210 bands. Therefore, we developed an alternate energy by choosing another matching term that approximates the two terms (2.14) and (2.13) from above.

If we choose the wavelet matching weights equal, $c_0 = c_1 = c_2$, the above term becomes a least squares match to the wavelet fused image. Respecting the idea that the best information we have away from the edges is the low resolution image, we construct a new matching image as a combination of the low resolution and the wavelet fused images. Denoting the wavelet fused image for the n th band with W_n , we define the new matching image Z_n to be

$$(2.15) \quad Z_n = \exp\left(-\frac{d}{|\nabla M|^2}\right) \cdot W_n + \left(1 - \exp\left(-\frac{d}{|\nabla M|^2}\right)\right) \cdot \uparrow H_n.$$

We then force our result to be close to the fused image Z_n by adding the term

$$(2.16) \quad E_a = \nu \sum_{n=1}^N \int_{\Omega} (u_n - Z_n)^2 dx$$

to our energy.

2.4. Spectral correlation preserving term. The previous terms constrain the colors within each band and introduce spatial information from the master image. The main focus was the spatial enhancement of the spectral image. However, the main use of hyperspectral images is the possibility of identifying materials based on the spectral information in the image. For this application it is common to look at the spectral z -direction at each pixel (see Figure 3). Most classification methods are not so much based on the magnitude but on the shape of the spectral signature, i.e., the signal in z -direction, at each pixel. Hence, it is crucial to preserve this signature in the sharpening process, which means we have to introduce a coupling term that preserves the relation between the bands pixel by pixel.

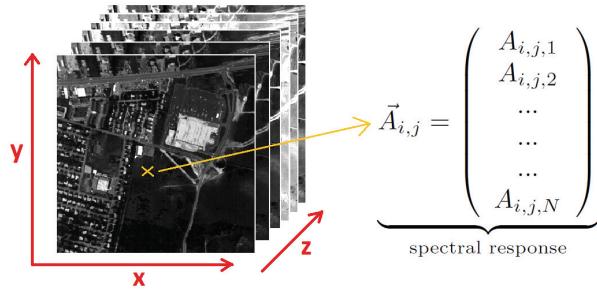


Figure 3. Common viewpoint for spectral response analysis.

To sharpen the image we have to allow the magnitude of the spectral signature to change. However, we want to preserve its shape. Instead of constraining the spectral signature itself, we propose keeping every possible ratio of two different spectral bands of our sharpened image equal to the ratio of the same bands of the original hyperspectral image. Mathematically, this means that we would like to obtain $\frac{u_i}{u_j} = \frac{\uparrow H_i}{\uparrow H_j} \Rightarrow u_i \cdot \uparrow H_j - u_j \cdot \uparrow H_i = 0$ at every pixel. We add the sum of the squares of the corresponding L_2 norms to our energy functional:

$$(2.17) \quad E_s = \mu \sum_{i,j=1, i < j}^N \int_{\Omega} (u_i \cdot \uparrow H_j - u_j \cdot \uparrow H_i)^2 dx.$$

For any spectral signature \vec{a} in the low resolution image and a corresponding signature \vec{b} in the high resolution image, the above energy term E_s reaches its minimum for $a(i) \cdot b(j) - a(j) \cdot b(i) = 0 \forall i, j$. For $b_j \neq 0$ we can write this as $a(i) = \frac{a(j)}{b(j)} \cdot b(i) = 0 \forall i, j$ and see that $\vec{a} \parallel \vec{b}$. The spectral angle, $SAM = \arccos\left(\frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\| \cdot \|\vec{b}\|}\right)$, is a widely used metric with which to compare and evaluate spectral signatures [50, 22, 4]. Notice that the spectral correlation preserving term E_s tries to keep the spectral vectors parallel and therefore minimizes the spectral angle. In

the minimum of (2.17) all classification methods based on the spectral angle mapper (SAM) will give the same result on the sharpened image as they do on the original spectral image.

We constructed two different energies. These are the VWP energy combining wavelet and spatial terms,

$$(2.18) \quad E_{VWP}(u) = E_g + E_s + E_w + E_c,$$

and the alternate energy (AVWP), which can be minimized in the spatial domain only. Putting the terms (2.6), (2.16), and (2.17) together, the AVWP energy functional is

$$(2.19) \quad E_{AVWP}(u) = E_g + E_s + E_a.$$

In this paper, we will focus on the results produced by the latter energy. We can show that there exists a minimizer for this energy functional under the smoothness assumption that $\text{div}(\theta) \in L^2$. The idea is that for $\text{div}(\theta) \in L^2$ we can complete the square by combining the geometry enforcing term E_g with the matching term E_a to get

$$\begin{aligned} E_g + E_a &= \sum_{n=1}^N \left[\gamma \int_{\Omega} |\nabla u_n| dx + \eta \int_{\Omega} \text{div}(\theta) \cdot u_n dx \right] + \nu \sum_{n=1}^N \int_{\Omega} (u_n - Z_n)^2 dx \\ &= \sum_{n=1}^N \left[\gamma \int_{\Omega} |\nabla u_n| + \nu \left(u_n - Z_n + \frac{\eta}{2\nu} \text{div}(\theta) \right)^2 - \left(\frac{\eta}{2\nu} \text{div}(\theta) \right)^2 + \nu Z_n^2 \right] dx. \end{aligned}$$

Since we can leave out the parts independent of u_n , this step converts our energy to a typical energy with an L^2 matching term and a TV regularization term. The existence and uniqueness proofs for such a functional can be found in [12]. For further details we refer the reader to the Master's thesis [34].

For VWP as well as for AVWP a couple of parameters need to be chosen. For both methods we need the parameters γ and η from the geometry introducing term and μ from the spectral correlation preserving term. Additionally, we have the matching parameters ν for AVWP and ν , c_0 , c_1 , and c_2 for VWP. As mentioned before, the parameters have to be adjusted depending on the type of image one wishes to produce. A higher weight on the spectral correlation preserving term will give better spectral quality in terms of the spectral angle but restricts the changes the algorithm is allowed to make to increase the spatial quality. Large matching parameters will typically result in an image that is closer to the low resolution image. For VWP we particularly have the choice of introducing more high frequency information from the master image in the wavelet domain. A big influence is a choice of geometry introducing parameters with $\gamma \neq \eta$, as we will demonstrate in the results section. Although the choice of parameters might seem difficult at first, all parameters have only to be tuned once for a certain type of data such that the parameter tuning is not too laborious.

To practically minimize (2.19) we compare two approaches. First we use gradient descent and discretize the resulting PDE semi-implicitly, similarly to the ideas for alternating direction minimization (ADI) introduced by Douglas, Peaceman, and Rachford [39, 20]. The second approach uses the more recent technique of “split Bregman” proposed by Goldstein and Osher [25]. Minimization by splitting techniques and Bregman iteration have gained a lot of attention

during the last year due to their superior behavior in minimizing singular energies. For linear constraints (like we will obtain in our minimization scheme) split Bregman has been proven to be equivalent to the augmented Lagrangian method. For an overview of different splitting methods and their connection, we refer the reader to [23].

3. Numerical minimization.

3.1. ADI minimization. The ADI approach has been used in image processing and for minimizing TV related functionals before (e.g., [11, 29, 10]). We make use of the necessary condition for minima that the first variation of the energy functional must equal zero. Unfortunately, the TV part of the energy is not differentiable, so we need to use subdifferential calculus or an additional smoothing similar to the ϵ -regularization we used in the calculation of θ_ϵ . We replace $|\nabla u_n|$ by $|\nabla u_n|_\epsilon$ and obtain for the n th band

$$(3.1) \quad \frac{\delta E_g}{\delta u_n} = - \left(\gamma \operatorname{div} \left(\frac{\nabla u_n}{|\nabla u_n|_\epsilon} \right) - \eta \operatorname{div}(\theta_\epsilon) \right),$$

$$(3.2) \quad \frac{\delta E_s}{\delta u_n} = 2\mu \sum_{j=1, j \neq n}^N (u_n \cdot \uparrow H_j - u_j \cdot \uparrow H_n) \uparrow H_j,$$

$$(3.3) \quad \frac{\delta E_a}{\delta u_n} = 2\nu(u_n - Z_n).$$

The optimality condition

$$(3.4) \quad \frac{\delta E_g}{\delta u_n} + \frac{\delta E_s}{\delta u_n} + \frac{\delta E_a}{\delta u_n} = 0$$

can be solved numerically by introducing an artificial time variable and solving

$$(3.5) \quad \frac{d}{dt} u_n = - \left(\frac{\delta E_g}{\delta u_n} + \frac{\delta E_s}{\delta u_n} + \frac{\delta E_a}{\delta u_n} \right)$$

to steady state. In our implementation we discretize in time implicitly for $\frac{\delta E_s}{\delta u_n}$ and $\frac{\delta E_a}{\delta u_n}$. For $\frac{\delta E_g}{\delta u_n}$ we lag the denominator by one time step (for not having to solve a nonlinear equation each time step) and then follow the ideas of ADI. We first take a step implicitly for the derivative in the x -direction and explicitly for the derivative in the y -direction and then change. This leaves us with a linear system where the corresponding matrix is tridiagonal such that each step can be solved extremely quickly, e.g., by the Thomas (TDMA) algorithm [18].

3.2. Split Bregman. The split Bregman method was recently proposed by Goldstein and Osher and is a fast minimization scheme for TV- or generally L_1 -related energies. Before the calculation of the Euler–Lagrange equation, a new variable is introduced to split the non-differentiable TV part from the rest of the energy. This way no additional ϵ -regularization as for the ADI method is needed. More specifically, minimizing (2.19) is equivalent to

$$(3.6) \quad \min_{u_n, d_n} E(u_n, d_n) = \min_{u_n, d_n} \gamma |d_n|_{L_1} + R(u_n) \quad \text{such that } d_n = \nabla u_n,$$

where $R(u_n)$ denotes the energy functional without the TV term:

$$(3.7) \quad \begin{aligned} R(u_n) = & \nu \int_{\Omega} (u_n - Z_n)^2 dx \\ & + \mu \sum_{j=1, j \neq n}^N \int_{\Omega} (u_n \cdot \uparrow H_j - u_j \cdot \uparrow H_n)^2 dx \\ & + \eta \int_{\Omega} \operatorname{div}(\theta) \cdot u_n dx. \end{aligned}$$

Applying the Bregman iteration, we obtain

$$(3.8) \quad \begin{aligned} (u_n^{k+1}, d_n^{k+1}) = & \underset{u_n, d_n}{\operatorname{argmin}} D_E^p(u_n, u_n^k, d_n, d_n^k) + \frac{\lambda}{2} \|d_n - \nabla u_n\|_2^2 \\ = & \underset{u_n, d_n}{\operatorname{argmin}} E(u_n, d_n) - \langle p_u^k, u_n - u_n^k \rangle - \langle p_d^k, d_n - d_n^k \rangle + \frac{\lambda}{2} \|d_n - \nabla u_n\|_2^2, \\ p_u^{k+1} = & p_u^k - \lambda \nabla^T (\nabla u_n^{k+1} - d_n^{k+1}), \\ p_d^{k+1} = & p_d^k - \lambda (d_n^{k+1} - \nabla u_n^{k+1}). \end{aligned}$$

This rather complicated formulation can be carried out in a simple two-step procedure without having to use the subgradients p_u and p_d . The scheme (3.8) is equivalent to

$$(3.9) \quad \begin{aligned} (u_n^{k+1}, d_n^{k+1}) = & \underset{u_n, d_n}{\operatorname{argmin}} \gamma \sqrt{d_x^2 + d_y^2} |_{L_1} + R(u_n) + \frac{\lambda}{2} \|d_n - \nabla u_n - b_n^k\|_2^2, \\ b_n^{k+1} = & b_n^k + (\nabla u_n^{k+1} - d_n^{k+1}). \end{aligned}$$

The optimization for u_n^{k+1} and d_n^{k+1} is done in an alternating way. Each part of the minimization is rather easy because (3.9) is differentiable in u_n , and the resulting optimality equation

$$(3.10) \quad \begin{aligned} & \left(2\nu + 2\mu \sum_{j=1, j \neq q}^N (\uparrow H_j)^2 - \lambda \Delta \right) u_q^{k+1} \\ & = 2\nu Z_q + \eta \operatorname{div}(\theta) + 2\mu \uparrow H_q \left(\sum_{j=1, j \neq q}^N u_j^{k*} \uparrow H_j \right) - \lambda \operatorname{div}(d_q^k - b_q^k) \end{aligned}$$

can approximately be solved by one sweep of Gauss–Seidel. Notice that we decoupled the bands by using $u_j^{k*} = u_j^{k+1}$ for $j < q$ and $u_j^{k*} = u_j^k$ for $j > q$ on the right-hand side of the equation. The solution for d_q^{k+1} for a given u_q^{k+1} can be given explicitly using the generalized shrinkage formula

$$(3.11) \quad (d_q^{k+1})_x = \max \left(s^k - \frac{\gamma}{\lambda}, 0 \right) \frac{\nabla_x u_q^{k+1} + (b_q^k)_x}{s^k},$$

$$(3.12) \quad (d_q^{k+1})_y = \max \left(s^k - \frac{\gamma}{\lambda}, 0 \right) \frac{\nabla_y u_q^{k+1} + (b_q^k)_y}{s^k},$$

with $s^k = \sqrt{|\nabla_x u_q^{k+1} + (b_q^k)_x|^2 + |\nabla_y u_q^{k+1} + (b_q^k)_y|^2}$. For more details on the method and its derivation we refer the reader to [25].

3.3. Convergence speed of split Bregman minimization. In addition to the results in [25], we want to give the convergence speed of this method before taking the alternating minimization between u_n^{k+1} and d_n^{k+1} into account. For a convergence analysis of the alternating minimization we refer the reader to [41].

Our observation holds not only for the application of split Bregman to anisotropic TV-regularization (as in our case) but for split Bregman methods in general, which is why we examine problems of the form

$$(3.13) \quad \min_u |\phi(u)|_{\ell^1} + R(u, f),$$

with a quadratic R and a convex regularization term $|\phi(u)|_{\ell^1}$. In our case we have $\phi(u) = \nabla u$ and $R(u, f)$ being the rest of the energy functional without the TV-regularization term. As seen above, the idea of split Bregman is to introduce a new variable $d = \phi(u)$ and solve the resulting constrained problem

$$(3.14) \quad \min_{u,d} J(u, d) = \min_{u,d} |d|_{\ell^1} + R(u, f) \text{ such that } d = \phi(u)$$

with Bregman iteration. To be able to give a convergence rate, we need the definition of a source condition.

Definition 3.1 (source condition). We denote

$$(3.15) \quad L(u, d, b) = J(u, d) - \langle b, d - \phi(u) \rangle.$$

We say that (\tilde{u}, \tilde{d}) satisfies a source condition if there exists a Lagrange multiplier \tilde{b} such that

$$(3.16) \quad L(\tilde{u}, \tilde{d}, b) \leq L(\tilde{u}, \tilde{d}, \tilde{b}) \leq L(u, d, \tilde{b}) \quad \forall u, d, b.$$

We can now state the convergence result.

Theorem 3.2. Let (\tilde{u}, \tilde{d}) be the solution to the problem $\phi(u) = d$. Then (\tilde{u}, \tilde{d}) satisfies a source condition, and the estimate

$$(3.17) \quad D_J^{p^{u_k}, p^{d_k}}((\tilde{u}, \tilde{d}), (u^k, d^k)) \leq \frac{\|\tilde{b}\|}{2\lambda k} = O\left(\frac{1}{k}\right)$$

holds, where \tilde{b} is the Lagrange multiplier from the source condition.

The convergence estimate is the result from Theorem 4.1 in [8] for Bregman iteration in general, where the theorem assumes that a source condition is satisfied. Hence, we immediately get a convergence rate for (u^k, d^k) toward a solution of the original minimization problem as soon as we can obtain a source condition of the true solution. Our contribution is to show that this is always the case for the split Bregman algorithm.

Lemma 3.3. The true solution \tilde{u} of the minimization problem (3.14) always satisfies the source condition (3.16).

Proof. The existence of a Lagrange multiplier in our case means that for

$$(3.18) \quad L(u, d, b) := |d|_{\ell^1} + R(u, f) - \langle b, d - \phi(u) \rangle$$

there exists a \tilde{b} such that

$$(3.19) \quad L(\tilde{u}, \tilde{d}, \tilde{b}) \leq L(u, d, \tilde{b}) \quad \forall u, d.$$

In other words (\tilde{u}, \tilde{d}) has to be the minimizer of the convex functional $L(\cdot, \cdot, \tilde{b})$. That means that a source condition is satisfied if and only if there exists a \tilde{b} such that the following two conditions hold.

1. Optimality condition for \tilde{d} :

$$(3.20) \quad p^{\tilde{d}} - \tilde{b} = 0, \quad p^{\tilde{d}} \in \partial |d|_{\ell^1}.$$

2. Optimality condition for \tilde{u} :

$$(3.21) \quad \tilde{q} + \phi^*(\tilde{b}) = 0, \quad \tilde{q} \in \partial R(\tilde{u}, f).$$

The true solution \tilde{u} with $\tilde{d} = \phi(\tilde{u})$ was a solution to the original minimization problem

$$(3.22) \quad \tilde{u} = \arg \min |d|_1 + R(u, f),$$

which means that it satisfies the optimality condition

$$(3.23) \quad \tilde{q} + \tilde{p} = 0$$

for $\tilde{q} \in \partial R(\tilde{u}, f)$ and $\tilde{p} \in \partial_u |\phi(\tilde{u})|$. The latter can be written as $\tilde{p} = \phi^*(p^{\tilde{d}})$ for a $p^{\tilde{d}} \in \partial_d |\tilde{d}|_{\ell^1}$. Therefore, we define $\tilde{b} := p^{\tilde{d}}$ and immediately obtain the two optimality conditions (3.20) and (3.21), which proves the lemma. ■

Equation (3.20) also gives us an estimate on \tilde{b} . For a nonzero \tilde{d} we have

$$(3.24) \quad \tilde{b} = \frac{\tilde{d}}{|\tilde{d}|} = \text{sign}(\tilde{d}),$$

but even for nondifferentiable parts where \tilde{d} is zero, we still know that the absolute value of a $p^{\tilde{d}} \in \partial |\tilde{d}|_{\ell^1}$ is bounded by one. Therefore, we get

$$(3.25) \quad \tilde{b} \leq |\Omega|.$$

Note that this estimate could even be much better if we knew more about the structure of \tilde{d} . Since we integrate over a function that we expect to be similar to the sign of \tilde{d} , we can hope that a sparse \tilde{d} would give us a much better convergence rate. In our case of using split Bregman on TV-regularization we have $\tilde{d} = \nabla \tilde{u}$. Thus, a sparse \tilde{d} would correspond to an image with only very few nonzero gradients, e.g., a piecewise constant image. Generally we can conclude that the fewer nonzero gradients our true, sharpened image has, the faster the method will converge. This is of course reasonable since fewer edges need to be sharpened.

4. Numerical results.

4.1. Results on multispectral data. In this section we present fusion results on the Quickbird satellite data, which consists of a four band multispectral image with a spatial resolution of 2.4m as well as a panchromatic image with a resolution of 0.6m. The images were taken at exactly the same time and exactly the same position such that only translational registration needed to be done to align the two images.

To evaluate the spatial and spectral quality of the images produced by our method, we compare it to the standard methods IHS, PCA, Brovey, and stationary wavelet fusion and to the other variational method, P+XS. The comparison is based on seven different image quality metrics that measure spectral quality and one metric measuring spatial quality. The metrics we used are the relative dimensionless global error in synthesis (ERGAS) (e.g., used in [22, 4]), the spectral angle mapper (SAM) [50, 22, 4], spectral information divergence (SID) [14], universal image quality index (Q-average) [49], root mean squared error (RMSE), relative average spectral error (RASE) [16], and correlation coefficients (CC) as used in [47] as well as the spatial quality measure introduced in [51], which we will refer to as filtered correlation coefficients (FCC). Additionally, we calculate the average change in variance between the fused image and the reference image in percent, which we refer to as VarC, and which should be a measure of how much contrast and spatial information has been introduced into the image. While the previous quality metrics have clear reference values, the optimal value for VarC is more difficult to give. On the one hand, the variance of the fused image should be as close to a ground truth as possible, but on the other hand, the more information introduced into the image the better.

Our method is specifically designed to preserve the spectral information from the low resolution image. Unfortunately, we do not have a dataset with a multispectral image at different resolutions, such that we are missing a dataset with ground truth to compare our results with. To show that we preserve the low resolution image spectral characteristics, we first compare the sharpened images to the upsampled low resolution images.

Second, to simulate a low resolution dataset with ground truth, we subsampled our data and compared the fusion results with the original data. This procedure is common for comparing multispectral fusion algorithms (cf. [48, 4, 47]). However, subsampling the data always includes model assumptions, namely, the direct problem of producing the low resolution data given the high resolution data. This is information we do not have for true satellite data. If we had this information, we would try to solve the corresponding inverse problem. In the literature authors have started to investigate the satellite sensor properties by looking at their modulation transfer functions (e.g., [3]). However, we would like to present a general image fusion method without taking specific sensor information into account, which is why in our experiments we simply used bicubic interpolation for the subsampling.

For all numerical results we show three different results for our algorithm—the results for VWP, for AVWP, and for AVWP with an increased contrast by choosing $\eta = 1.3\gamma$ as discussed in section 2.2, which we will refer to as AVWP*. The following parameters were used: For AVWP* we used $\gamma = 1$, $\eta = 1.3$, $\mu = 50$, and $\nu = 4$ with a second level stationary wavelet decomposition using MATLAB’s “sym4” decomposition. For AVWP we used $\gamma = 1$, $\eta = 1$, $\mu = 50$, and $\nu = 4$ with a second level stationary wavelet decomposition using MATLAB’s

sym4 decomposition. For VWP we used $\gamma = 1$, $\eta = 1$, $\nu = 60$, and $\mu = 200$ with a second level stationary wavelet decomposition using MATLAB's sym4 and wavelet matching parameters $c_2 = 4$, $c_1 = 1.5$, and $c_0 = 0.5$.

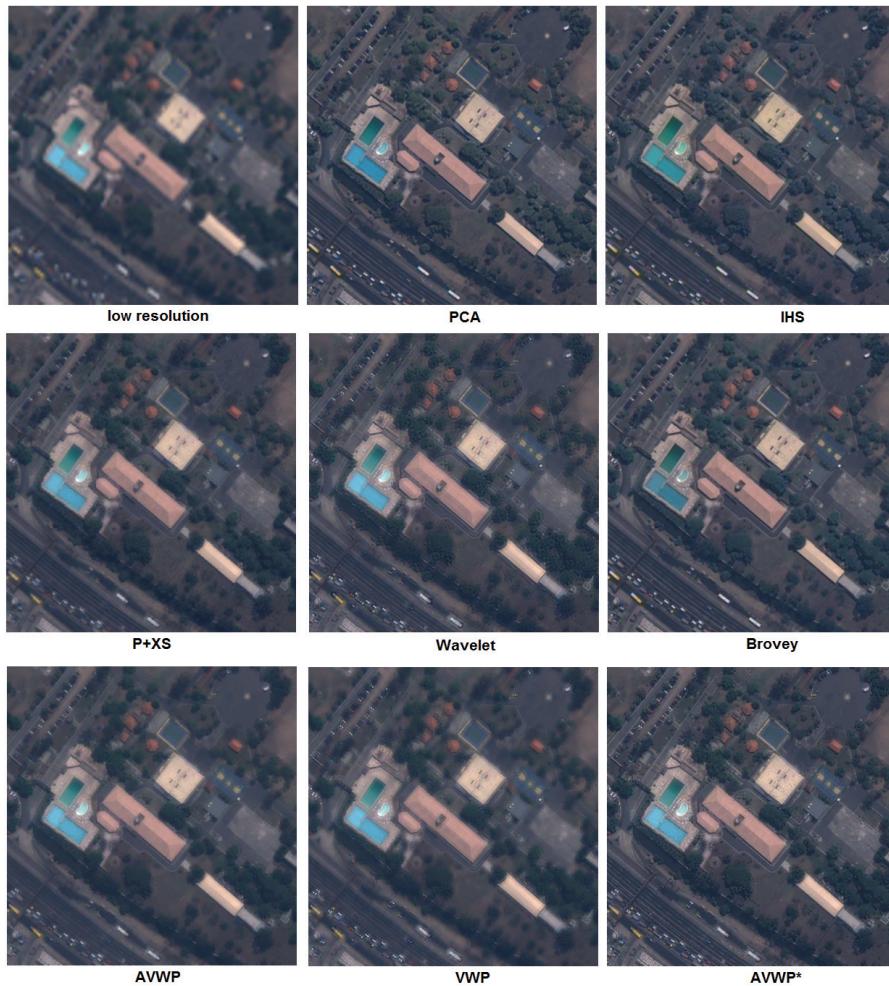


Figure 4. Results for sharpening multispectral images. Copyright DigitalGlobe, NextView license, 2003.

Figures 4 and 5 show the fusion results for different pan-sharpening methods as well as an upsampled low resolution image using bicubic interpolation. We can see that all methods greatly improve the visual quality of the low resolution image. The standard methods, PCA, IHS, and Brovey, always give visually satisfying results. However, by taking a close look at the color of the fused images and the colors of the low resolution image we can see that these methods tend to change some of the colors. Particularly, the colors of the trees and the color of the swimming pool in Figure 4 seem to be slightly off. This is a sign of spectral distortion. To quantify spectral quality we calculated the above-mentioned quality metrics with respect to the low resolution image. The results are shown in Tables 1 and 2, respectively.

The quality metrics indicate that PCA, IHS, and Brovey have the highest spatial quality,

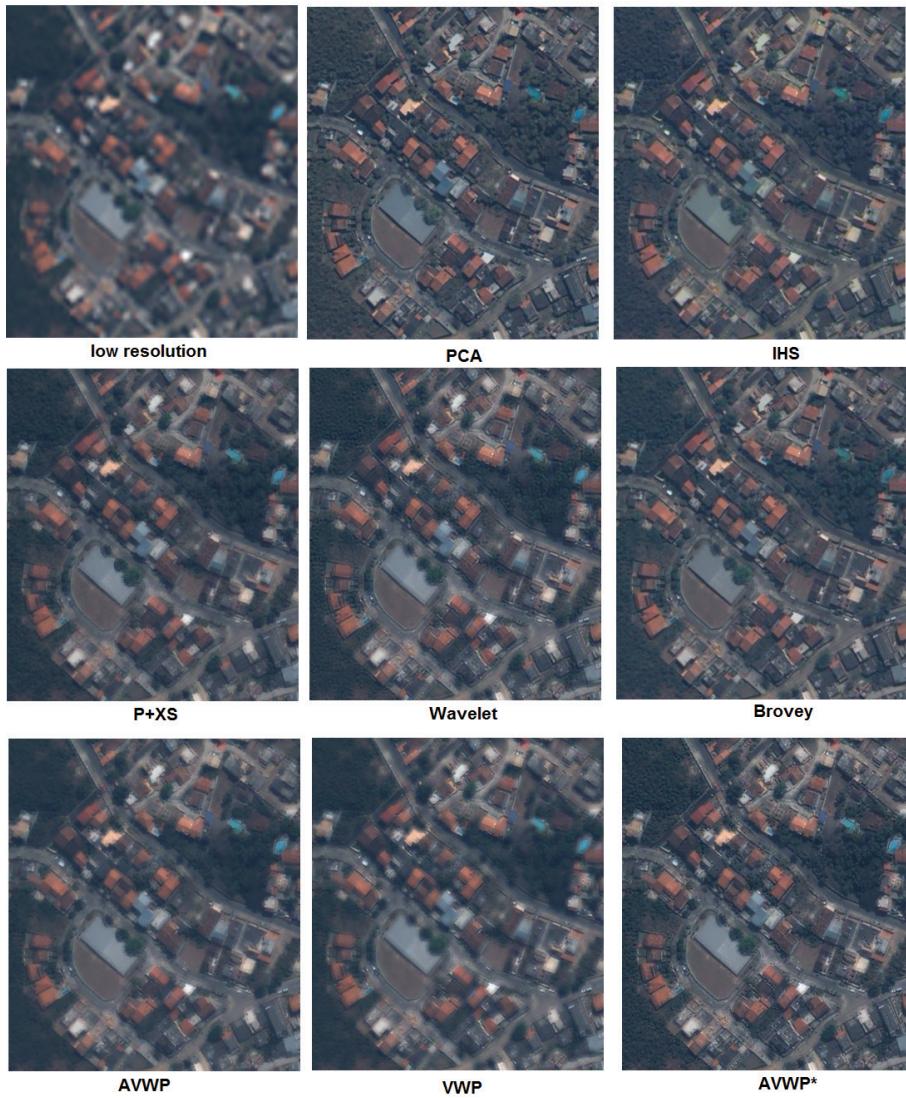


Figure 5. Results for sharpening multispectral images. Copyright DigitalGlobe, NextView license, 2003.

but they also distorted the spectral quality the most. The fact that Brovey always satisfies $SAM = 0$ comes from the fact that each band is multiplied by the same factor, and therefore SAM never changes. However, the other quality metrics show that Brovey still changes the spectral information.

We can see that the VWP and AVWP fusion results are not as sharp as the methods mentioned above. However, our goal was to sharpen the low resolution image while preserving most of the spectral information. Looking at the image quality metrics, we can see that AVWP and VWP fusion results are close to the low resolution image in terms of spectral information. For both images VWP and AVWP have the best values in ERGAS, RASE, RMSE, and CC and are second in SAM, SID, and QAVE. The AVWP* image has very sharp edges and high

Table 1Image quality metrics for Figure 4, SID^* in 10^{-3} , CC^{**} in 10^{-2} .

	ERGAS	QAVE	RASE	RMSE	SAM	SID*	CC**	FCC	VarC
Brovey	3.29	0.923	13.3	51.5	0	0	4.11	0.977	-3.67
IHS	3.35	0.917	12.7	49.2	1.24	5.54	5.14	0.982	-4.02
PCA	3.29	0.938	13	50.2	1.43	5.57	3.69	0.978	-4.36
Wavelet	1.96	0.976	7.9	30.5	0.72	0.46	1.48	0.973	4.9
P+XS	1.87	0.982	7.4	28.6	1.35	0.99	2.3	0.839	1.14
AVWP*	2.52	0.961	10.1	39.3	0.51	0.18	3.2	0.908	10.82
AVWP	1.6	0.984	6.4	25	0.48	0.15	1.29	0.907	2.76
VWP	1.59	0.984	6.4	24.6	0.22	0.05	1.08	0.901	4.4
lowres	0	1	0	0	0	0	0	0.523	0

Table 2Image quality metrics for Figure 5, SID^* in 10^{-3} , CC^{**} in 10^{-2} .

	ERGAS	QAVE	RASE	RMSE	SAM	SID*	CC**	FCC	VarC
Brovey	3.35	0.879	13.6	50	0	0	3.83	0.977	-0.13
IHS	3.49	0.864	13.2	48.5	1.28	1.17	6	0.979	-1.66
PCA	3.63	0.883	14	51.5	1.95	2.68	5.79	0.969	-3.89
Wavelet	2.18	0.954	8.7	32.1	0.74	0.49	3.13	0.973	8.98
P+XS	2.13	0.967	8.3	30.6	1.31	0.9	3.06	0.842	2.35
AVWP*	2.87	0.923	11.6	42.7	0.42	0.12	6.18	0.912	21.28
AVWP	1.75	0.969	7.1	26.1	0.42	0.11	2.16	0.902	5.01
VWP	1.64	0.973	6.6	24.2	0.23	0.05	2.33	0.897	7.42
lowres	0	1	0	0	0	0	0	0.505	0

visual quality, but we can see in the quality metrics that we gain the visual quality at the price of losing some of the spectral information of the low resolution image. Nevertheless, we still get much better values than we do with the standard methods. Particularly, the idea of minimizing a term to keep the spectral angle at zero still gives very good SAM and SID metric values.

The metric values for spatial quality are a little surprising. In our opinion the result that P+XS clearly has the worst spatial quality cannot be confirmed by visual analysis. Furthermore, we would argue that the visual quality of AVWP* is equal to the top methods, PCA, IHS, and Brovey. It seems to be difficult to find a metric that measures the “sharpness” of an image precisely. Moreover, almost any metric that we believe to measure sharpness could be used in a variational framework by finding the image that minimizes this particular metric. The difficulty in measuring spatial quality is underlined by the metric values for the change in variance, VarC. While IHS, PCA, and Brovey reduce the variance in comparison to the blurry low resolution image, all other methods increase it. The largest increase can be seen for the AVWP* method, indicating that many details have been introduced into the fused image.

As mentioned above, we also simulated fusion data with ground truth by subsampling the given data. Figures 6, 7, and 8 show the fusion results and Tables 3, 4, and 5 the corresponding quality metric values. Notice that the images denoted by “ground truth” are now the original

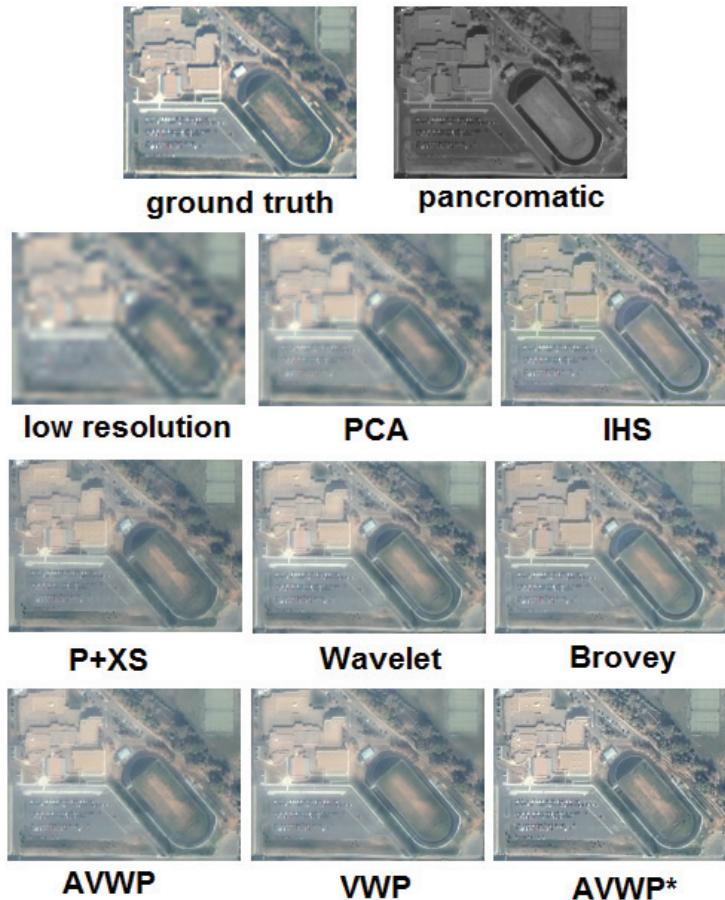


Figure 6. Results for sharpening multispectral images.

images and “low resolution” images are the images we obtained by first subsampling the ground truth and then upsampling again using bicubic interpolation for both operations. Looking at the quality metrics, we can see that although the difference between the methods seems to be smaller, VWP and AVWP still perform very well. The contrast increased AVWP images give much better quality values than in the previous comparison, which shows that the contrast improvement could be of practical relevance. Comparing to the ground truth, the spectral quality of AVWP* does not seem to be bad. Particularly interesting is that the variance of the fused images is still low in comparison to our simulated ground truth, which means that either the contrast was reduced or not enough high frequency details were injected. AVWP* has the highest variance among all methods and in image 8 even exceeds the variance of the ground truth.

The low resolution image seems to be far from the true data in terms of absolute pixel values, e.g., in the RMSE and RASE metrics. At least for the first test image all methods were able to improve these values in comparison to the low resolution image. However, the SAM value is generally best for the low resolution image, which encourages and strengthens

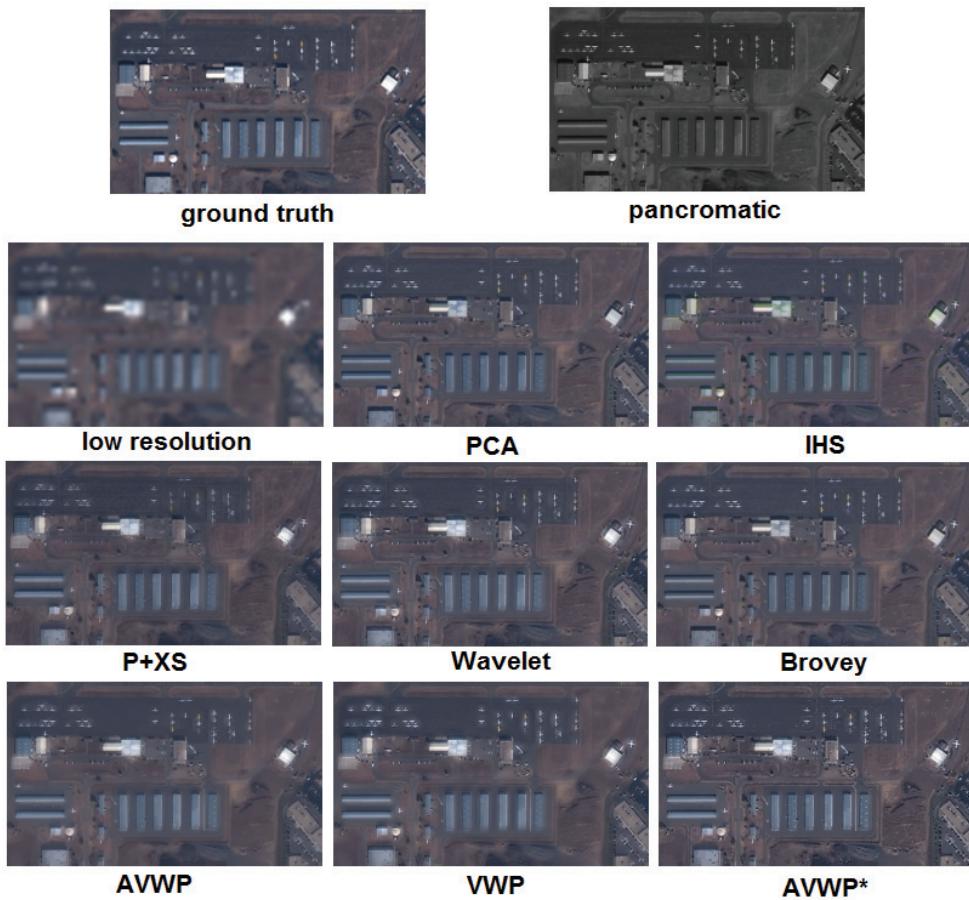


Figure 7. Results for sharpening multispectral images. Copyright DigitalGlobe, NextView license, 2003.

our idea to explicitly constrain this spectral metric in our energy functional. The SAM value of all VWP and AVWP fusion results is the best or close to the best. The combination of this result and the fact that VWP could be used to sharpen an arbitrary number of bands motivates the extension of VWP to hyperspectral image fusion.

4.2. Results on hyperspectral data. To demonstrate that our algorithm easily extends to higher dimensional images we worked with two datasets: an 82-band AVIRIS hyperspectral image of San Diego harbor [36] and a 210 band HYDICE hyperspectral image of an urban scene in Texas which is freely available online [46]. Since the datasets were not accompanied by a master image, we looked at the same scene on Google Maps [27] (which offers very high resolution images) and extracted several scenes as our master images. Unfortunately, the hyperspectral images are spatially distorted in comparison to the Google Maps images. For the examples shown in this report we did the required registration on small parts of the image manually. Results of sharpening hyperspectral images are shown in Figure 9.

In general we can say that the visual quality is greatly increased by the sharpening process. While in the low resolution hyperspectral images small objects can almost not be visually

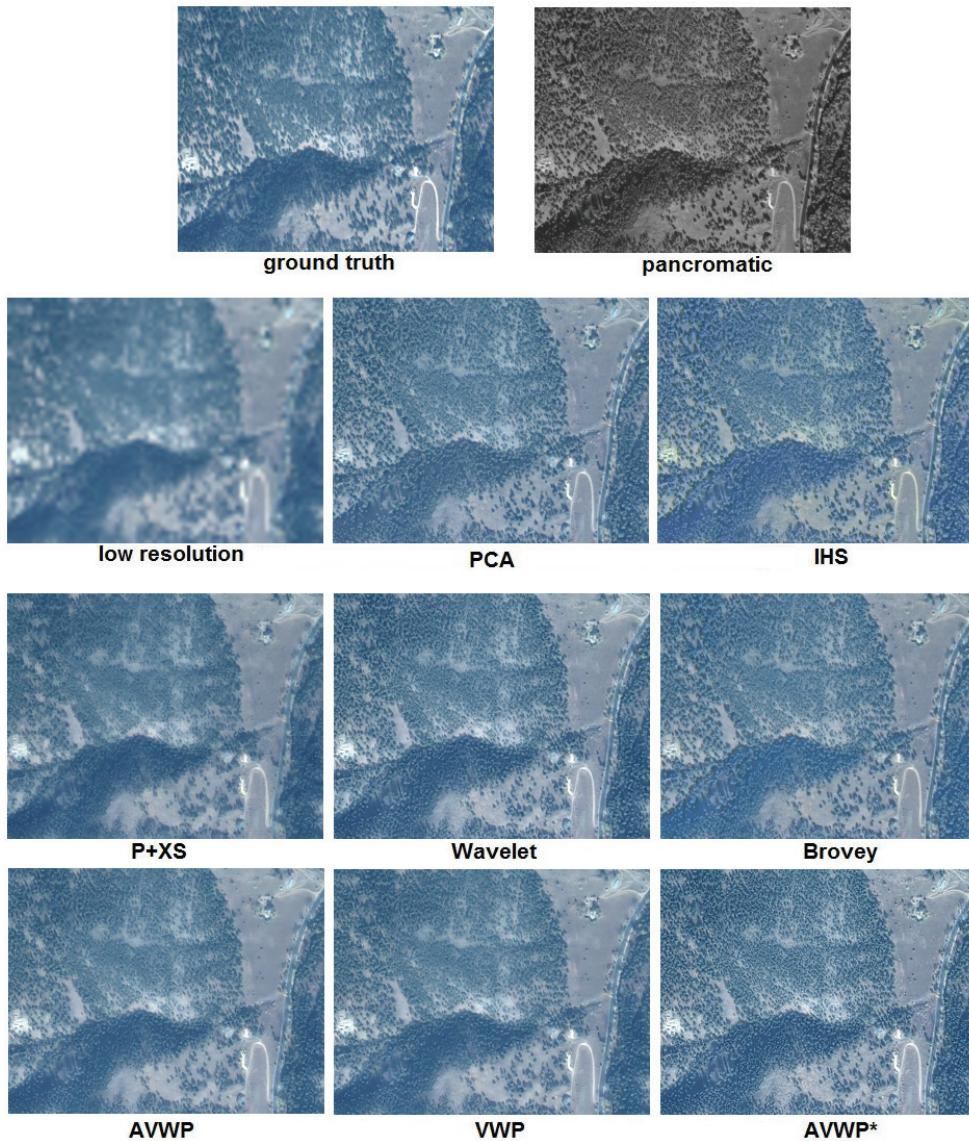


Figure 8. Results for sharpening multispectral images. Copyright DigitalGlobe, NextView license, 2003.

identified, the sharpened images are very close to the master images in terms of spatial quality. We should mention that we do not know when the Google Maps picture was taken. The boxes in the second image of Figure 9, for instance, could be things we see in the master image that were not present in the hyperspectral image and therefore appear as phantoms. One has to be careful with the introduced spatial information if the two images were not taken at the same time.

As mentioned earlier, the most important issue in sharpening hyperspectral images is not to change the spectral signature during the sharpening process since this information is used for material classification. To investigate how the spectral signature changes, we select two

Table 3Image quality metrics for Figure 6, SID^* in 10^{-3} , CC^{**} in 10^{-2} .

	ERGAS	QAVE	RASE	RMSE	SAM	SID*	CC**	FCC	VarC
Brovey	2.88	0.846	11.8	31.6	3.62	7.37	5.64	0.952	-45.38
IHS	3.05	0.823	12.3	33	4.03	9.2	6.94	0.951	-44.9
PCA	3.18	0.835	13.5	36.1	4.4	8.61	5.08	0.764	-38.05
Wavelet	2.93	0.871	12.5	33.5	3.72	7.68	2.19	0.944	-32.39
P+XS	3.63	0.848	14.2	38	3.55	6.85	5.5	0.794	-42.63
AVWP*	2.81	0.883	12	32	3.62	7.31	2.53	0.919	-23.7
AVWP	2.92	0.869	12.5	33.3	3.62	7.32	4.34	0.868	-35.7
VWP	2.89	0.875	12.4	33.2	3.64	7.42	3.27	0.915	-33.5
lowres	3.68	0.78	15.6	41.6	3.62	7.37	8.93	0.481	-41.79
Reference	0	1	0	0	0	0	0	1	0

Table 4Image quality metrics for Figure 7, SID^* in 10^{-3} , CC^{**} in 10^{-2} .

	ERGAS	QAVE	RASE	RMSE	SAM	SID*	CC**	FCC	VarC
Brovey	3.61	0.843	14.6	39.9	2.23	3.01	0.92	0.984	-37.02
IHS	3.69	0.834	14.6	40	2.52	3.74	3.3	0.985	-35.91
PCA	3.56	0.851	14.5	39.7	2.27	3.06	2.31	0.985	-37.18
Wavelet	3.2	0.889	13.1	35.7	2.42	3.48	1.36	0.983	-24.31
P+XS	3.47	0.889	13.6	37	2.47	3.37	0.98	0.863	-32.54
AVWP*	3.38	0.882	13.8	37.5	2.24	3.02	3.84	0.944	-13.45
AVWP	3.03	0.898	12.4	33.9	2.24	3.04	0.65	0.938	-28.76
VWP	2.93	0.907	12	32.8	2.27	3.1	0.83	0.943	-24.16
lowres	3.41	0.867	14	38.2	2.23	3.01	2.44	0.441	-35.64
Reference	0	1	0	0	0	0	0	1	0

Table 5Image quality metrics for Figure 8, SID^* in 10^{-3} , CC^{**} in 10^{-2} .

	ERGAS	QAVE	RASE	RMSE	SAM	SID*	CC**	FCC	VarC
Brovey	3.91	0.759	17.2	35.4	4.47	9.65	8.96	0.982	-29.01
IHS	4.27	0.717	17.9	36.8	5.27	13.68	8.77	0.981	-19.85
PCA	3.95	0.803	17.2	35.4	4.67	10.75	6.69	0.971	-36.18
Wavelet	3.78	0.826	16.7	34.5	4.68	10.74	9.32	0.986	-7.98
P+XS	3.87	0.851	16.9	34.8	4.49	9.74	7.79	0.881	-28.08
AVWP*	4.39	0.758	19	39.1	4.48	9.68	10.7	0.951	20.75
AVWP	3.57	0.833	16.2	33.4	4.48	9.69	8.1	0.959	-17.56
VWP	3.69	0.822	16.7	34.4	4.5	9.82	8.18	0.95	-16.35
lowres	3.74	0.83	17.8	36.6	4.47	9.65	6.69	0.334	-34.71
Reference	0	1	0	0	0	0	0	1	0

pixels in the 82 band hyperspectral scene and look at their spectral signature. Figure 10 shows the result of the sharpening process as well as the two pixels we selected to examine the change of spectral response during the sharpening process.

The upper left pixel is in the middle of a box where there are no edges or texture even in the

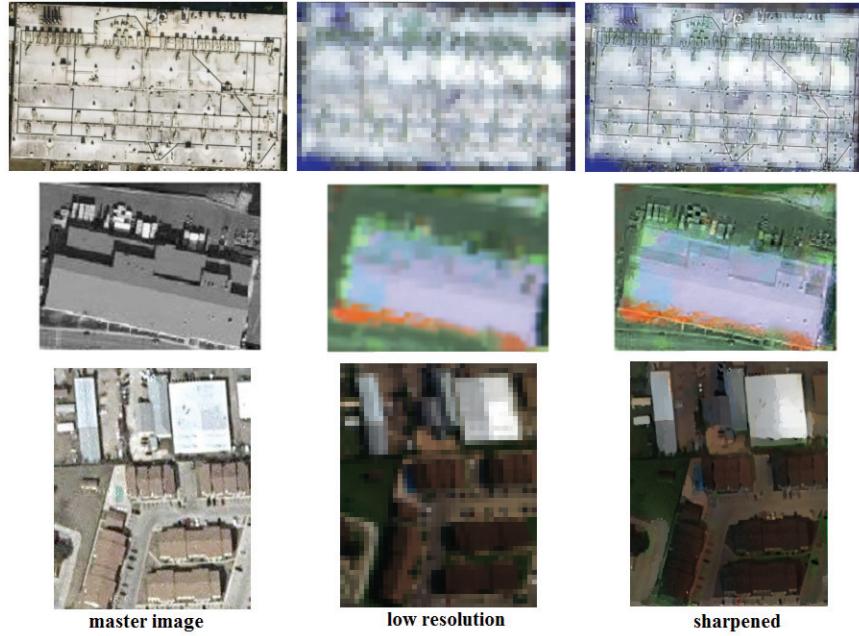


Figure 9. Results for sharpening hyperspectral images.

master image. At these pixels we match our current iteration to the low resolution image and therefore do not change the intensity in any band. This can be seen in its spectral signature before and after the image fusion. The curves match almost exactly, and the signature does not change. At the other pixel we are sure to change the signature because we want to increase the contrast and enhance the edge. Our energy model allows this increase of contrast but enforces the frequency vectors to stay parallel. Investigating the spectral responses, we can see that the signatures are different after the sharpening process. However, normalizing both signatures to 1 they again match almost perfectly. The spectral angle is kept at almost zero degrees such that any algorithm based on SAM will perform equally well on the sharpened image as on the low resolution data.

4.3. Comparison of split Bregman and ADI. For our energy minimization problem the most important difference between these methods is the additional regularization parameter ϵ that is needed by the ADI method to smooth the TV term. It can be shown that the time step for the ADI method is restricted proportional to ϵ . The larger we choose ϵ , the faster ADI will be, but, on the other hand, we minimize an energy that is different from the original energy. Edges will be smoothed out for too-large ϵ .

Split Bregman does not need the additional regularization and minimizes our energy more precisely. However, we still use an ϵ -regularization for the calculation of θ_ϵ . This effect might partly compensate for the significant advantages the split Bregman method seems to have for standard TV denoising, as shown in [25].

Our original data multispectral intensity values are in the range between 0 and a little above 1000. To preserve this precision during the sharpening process, the ϵ -regularization

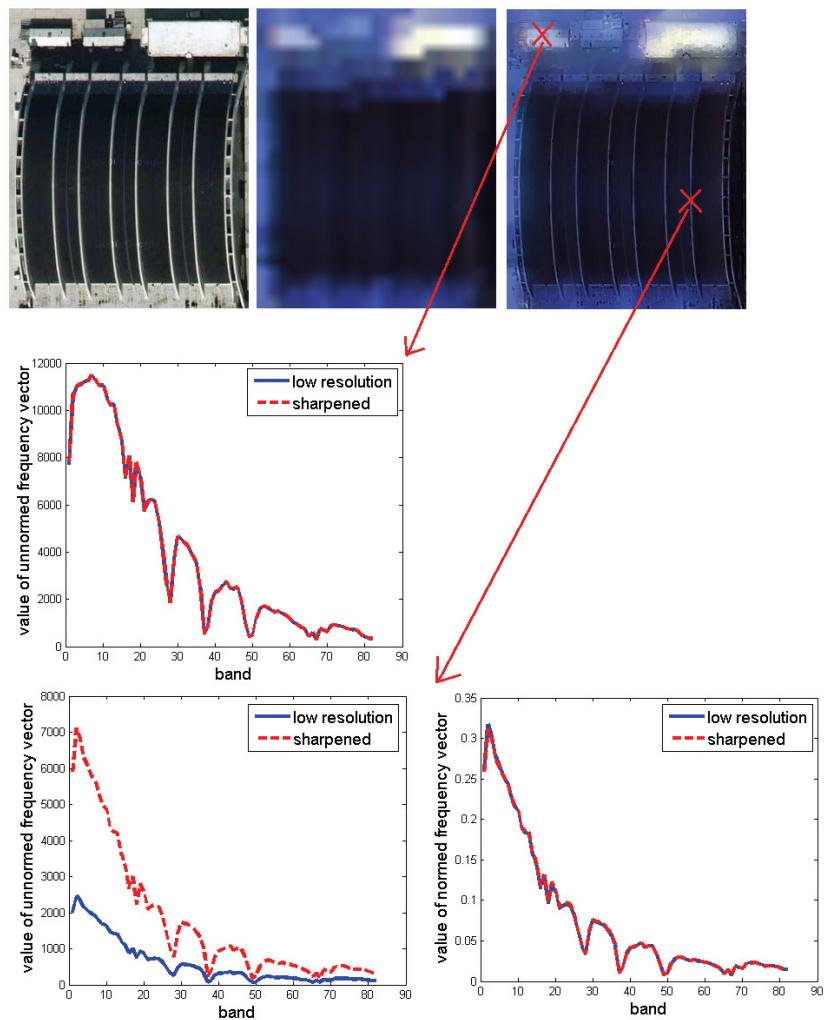


Figure 10. Behavior of the spectral signature during the sharpening process.

should not exceed the reciprocal of the range of values, i.e., $\epsilon \leq 10^{-3}$. Comparing the two minimization methods is difficult, since they minimize slightly different energies, the additionally regularized and the original energies. Figure 11 shows the decay of both types of energy for both methods. We can see that split Bregman minimizes the energy slightly faster. Note that ADI's runtime depends on the choice of ϵ , while split Bregman's runtime is independent. We can see that the final energy value for ADI is lower than the final split Bregman energy in the additionally regularized case and vice versa for the original energy.

To get a better and fairer comparison, we rewrite split Bregman such that it minimizes the additional regularized energy. However, this is just for comparison purposes, because it significantly slows down the method. Nevertheless, we can directly compare the two methods this way. To show the effect of an even smaller ϵ (which would, for instance, be necessary for higher resolved images), we ran the two methods for $\epsilon = 10^{-4}$.

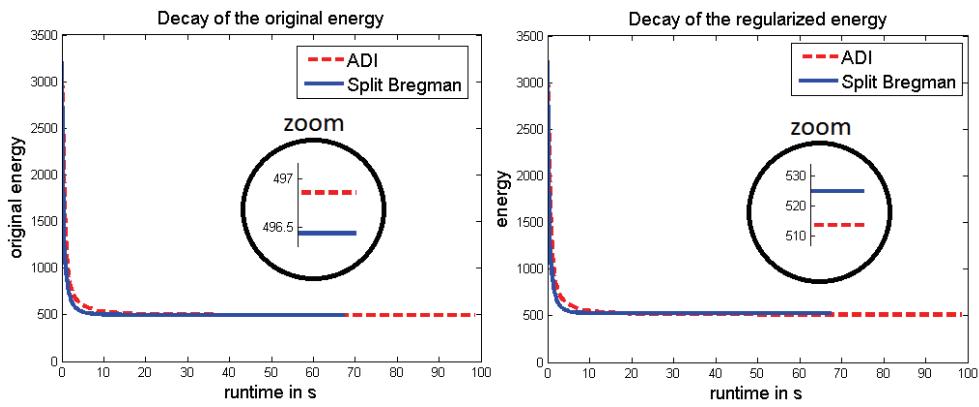


Figure 11. Comparison of ADI and split Bregman for energy minimization.

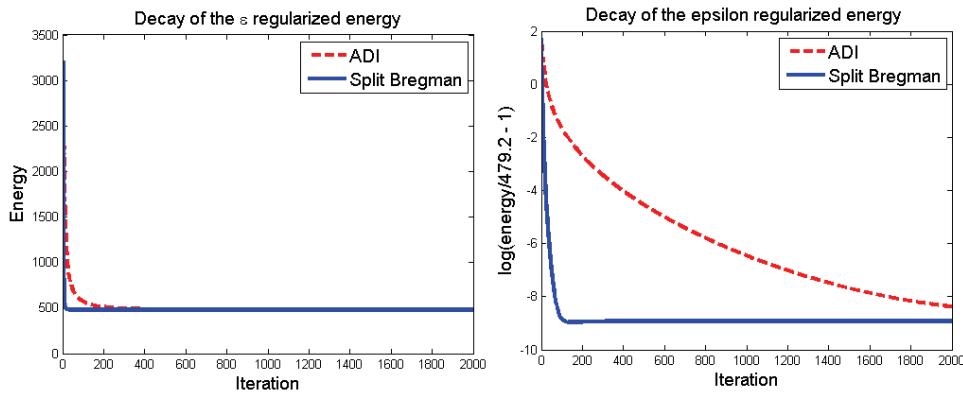


Figure 12. Comparison of ADI and split Bregman for energy minimization.

Figure 12 shows the decay of energy over 2000 iterations on the left. Both methods have a steep drop in energy at the beginning, but ADI slows down and takes many more iterations to reach the same energy as split Bregman. The minimum energy of split Bregman is 479.26, while the minimum energy of ADI is 479.31. The split Bregman energy minimum is not reached at the last iteration. At some point the energy does not decrease anymore; it slightly increases and then reaches an absolute steady state. This could be due to numerical errors, rounding errors, and the fact that we use alternating minimization for the bands and for the two split Bregman variables as well as the fact that we solve for the variables u_i only approximately by one step of a Gauss–Seidel algorithm. However, the difference to the minimal energy is 0.0026 , which amounts to an average change of less than $4 \cdot 10^{-8}$ per pixel. This precision is very hard to resolve and would require much more computational effort. We can guess that the true minimal energy might be around 479.2. To be able to better resolve the speed of decay, we use a semilogarithmic plot of the energy divided by the minimal energy minus 1 (see Figure 12, right).

Split Bregman minimizes the energy very quickly and reaches a minimum after about 150 iterations. ADI takes longer, particularly to reach a very high precision solution. However,

we should mention that this examination (particularly with $\epsilon = 10^{-4}$) is of interest for very exact solutions. It becomes more and more important the better the resolution of the sensors becomes. For our current data the choice $\epsilon = 10^{-3}$ (as examined above) probably is sufficient, and for visually satisfying results displayed in RGB colors with values in [0, 255] we could even choose $\epsilon = 5 * 10^{-3}$.

Notice that variational methods like VWP, AVWP, and P+XS are generally much more computationally expensive than simple explicit methods like IHS and Brovey. However, as we can see in the left part of Figure 11, the split Bregman algorithm needs only about 20 seconds to compute the sharpened image for an image size of 330 by 280 pixels on a laptop computer with 3GB memory and a dual core 2GHz processor. Furthermore, and even more important, the simple methods like IHS and Brovey are based on the assumption of a linear relation between the master image and the spectral bands such that they do not extend to hyperspectral imaging as VWP does.

5. Conclusions. We proposed a variational method for sharpening high dimensional spectral images which extends from usual pan-sharpening to an arbitrary number of bands with a master image that does not need to be panchromatic. The energy model incorporates the alignment of all unit normal vectors of the level sets of each band with the master image and uses a fidelity term with a combination of the original and a wavelet-fused spectral image. By keeping the ratio of all bands constant, we ensure spectral quality. Large weights on this spectral quality term force the spectral angle between the original and the sharpened spectral images to be close to zero. For higher visual quality we chose different weights in the geometry forcing term to introduce higher edge curvature and therefore higher contrast.

For the optimization we compared a gradient descent ADI method with the split Bregman algorithm and can conclude that split Bregman seems to be the superior minimization method. A theoretical convergence speed for split Bregman was proved.

For future research one could try to incorporate technical information about the satellite sensors into the variational framework, and other types of sensors and images could be used. To automate the sharpening process a robust registration method is needed. To reduce the bleeding of colors over some of the edges, deblurring could be included in the sharpening method. Furthermore, this method could be combined with hyperspectral analysis methods such as demixing to improve not only the spatial resolution but also the spectral resolution of the image. New detection and classification methods especially suitable for sharpened images which take the spectral as well as the spatial information into account could be developed.

Acknowledgments. The authors would like to thank Melissa Strait, Sheida Rahmani, and Daria Merkurev for their help and especially for providing the MATLAB code for all image quality metrics. We would also like to thank Julia Dobrosotskaya, Stanley Osher, Jerome Darbon, and Ernie Esser for their advice.

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. Geosci. Remote Sensing, 40 (2002), pp. 2300–2312.

- [2] B. AIAZZI, L. ALPERONE, S. BARONTI, I. PIPPI, AND M. SELVA, *Generalized Laplacian pyramid-based fusion of MS+P image data with spectral distortion minimization*, in ISPRS Commission III Symposium 2002, Graz, Austria, 2002.
- [3] L. ALPARONE, B. AIAZZI, S. BARONTI, AND A. GARZELLI, *A global quality measurement of pan-sharpened multispectral imagery*, in Proceedings of the 2003 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Washington, DC, 2003, pp. 458–460.
- [4] L. ALPARONE, L. WALD, J. CHANUSSOT, C. THOMAS, P. GAMBA, AND L. M. BRUCE, *Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest*, IEEE Trans. Geosci. Remote Sensing, 45 (2007), pp. 3012–3021.
- [5] C. BALLESTER, V. CASELLES, L. IGUAL, AND J. VERDERA, *A variational model for P+XS image fusion*, Internat. J. Comput. Vision, 69 (2006), pp. 43–58.
- [6] L. M. BREGMAN, *The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming*, USSR Comput. Math. Math. Phys., 7 (1967), pp. 200–217.
- [7] M. BURGER AND S. OSHER, *Convergence rates of convex variational regularization*, Inverse Problems, 20 (2004), pp. 1411–1421.
- [8] M. BURGER, E. RESMERITA, AND L. HE, *Error estimation for Bregman iterations and inverse scale space methods in image restoration*, Computing, 81 (2007), pp. 109–135.
- [9] J.-F. CAI, S. OSHER, AND Z. SHEN, *Fast Linearized Bregman Iteration for Compressed Sensing and Sparse Denoising*, UCLA CAM Rep. 08–37, UCLA, Los Angeles, CA, 2008.
- [10] D. CALVETTI AND L. REICHEL, *Application of ADI iterative methods to the restoration of noisy images*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 165–186.
- [11] Y. CHA AND S. KIM, *Edge-forming methods for image zooming*, J. Math. Imaging Vision, 25 (2006), pp. 353–364.
- [12] A. CHAMBOLLE AND P.-L. LIONS, *Image recovery via total variation minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.
- [13] T. F. CHAN, J. SHEN, AND H.-M. ZHOU, *A total variation wavelet inpainting model with multilevel fitting parameters*, Proc. SPIE, 6313 (2006), p. 63130C.
- [14] C.-I. CHANG, *Spectral information divergence for hyperspectral image analysis*, in Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IEEE, Washington, DC, 1999, pp. 509–511.
- [15] Y. CHIBANI AND A. HOUACINE, *Redundant versus orthogonal wavelet decomposition for multisensor image fusion*, Pattern Recognition, 36 (2003), pp. 879–887.
- [16] M. CHOI, *A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter*, IEEE Trans. Geosci. Remote Sensing, 44 (2006), pp. 1672–1682.
- [17] M. CHOI, H.-C. KIM, N. I. CHO, AND H. O. KIM, *An improved intensity-hue-saturation method for IKONOS image fusion*, Int. J. Remote Sensing, submitted; also available online from <http://amath.kaist.ac.kr/research/paper/06-9.pdf>.
- [18] S. CONTE AND C. DEBOOR, *Elementary Numerical Analysis*, McGraw–Hill, New York, 1972.
- [19] J. DOBROSOTSKAYA AND A. L. BERTOZZI, *A wavelet-Laplace variational technique for image deconvolution and inpainting*, IEEE Trans. Image Process., 17 (2008), pp. 657–663.
- [20] J. DOUGLAS AND D. W. PEACEMAN, *Numerical solution of two-dimensional heat-flow problems*, AIChE J., 1 (1955), pp. 505–512.
- [21] Q. DU, O. GUNGOR, AND J. SHAN, *Performance evaluation for pan-sharpening techniques*, in Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IEEE, Washington, DC, 2005, pp. 4264–4266.
- [22] Q. DU, N. H. YOUNAN, R. KING, AND V. P. SHAH, *On the performance evaluation of pan-sharpening techniques*, IEEE Geosci. Remote Sensing Lett., 4 (2007), pp. 518–522.
- [23] E. ESSER, *Primal Dual Algorithms for Convex Models and Applications to Image Restoration, Registration and Nonlocal Inpainting*, Tech. report, Ph.D. thesis, UCLA CAM Report 10-31, UCLA, Los Angeles, CA, 2010.
- [24] A. GARZELLI AND F. NENCINI, *Fusion of panchromatic and multispectral images by genetic algorithms*, in Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IEEE, Washington, DC, 2006, pp. 3810–3813.

- [25] T. GOLDSTEIN AND S. OSHER, *The split Bregman method for L1 regularized problems*, SIAM J. Imaging Sci., 2 (2009), pp. 323–343.
- [26] M. GONZALEZ-AUDICANA, X. OTAZU, O. FORS, AND J. ALVAREZ-MOZOZ, *A low computational-cost method to fuse IKONOS images using the spectral response function of its sensors*, IEEE Trans. Geosci. Remote Sensing, 44 (2006), pp. 1683–1691.
- [27] GOOGLE, INC., *Google Maps*, <http://maps.google.com>.
- [28] G. HONG AND Y. ZHANG, *The effects of different types of wavelets on image fusion*, in Proceedings of the International Conference on Image Processing, IEEE, Washington, DC, 1995, pp. 248–251.
- [29] K. JOO AND S. KIM, *PDE-based Image Restoration, II: Numerical Schemes and Color Image Denoising*, Tech. report 2003-08, Department of Mathematics, University of Kentucky, Lexington, KY, 2003.
- [30] R. L. KING AND J. WANG, *A wavelet based algorithm for pan-sharpening Landsat 7 imagery*, in Proceedings of the IEEE Geoscience and Remote Sensing Symposium, IEEE, Washington, DC, 2001, pp. 849–851.
- [31] O. M. LYSAKER, S. OSHER, AND X.-C. TAI, *Noise removal using smoothed normals and surface fitting*, IEEE Trans. Image Process., 13 (2004), pp. 1345–1357.
- [32] F. MALGOUYRES, *Mathematical analysis of a model which combines total variation and wavelet for image restoration*, J. Inform. Process., 2 (2002), pp. 1–10.
- [33] N. MEMARSADEGH, J. L. MOIGNE, AND D. MOUNT, *Image fusion using cokriging*, in Proceedings of the IEEE Geoscience and Remote Sensing Symposium, IEEE, Washington, DC, 2006, pp. 2518–2521.
- [34] M. MÖLLER, *A Variational Approach to Sharpening High Dimensional Images*, Master's thesis, Westfälische Wilhelms-Universitaet, Muenster, Germany, 2009.
- [35] M. MÖLLER, T. WITTMAN, AND A. L. BERTOZZI, *A variational approach to hyperspectral image fusion*, in Proceedings of the SPIE Conference on Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV, SPIE, Bellingham, WA, 2009.
- [36] NASA, *NASA Jet Propulsion Laboratory*, <http://aviris.jpl.nasa.gov/>.
- [37] S. OSHER, M. BURGER, D. GOLDFARB, J. XU, AND W. YIN, *An iterative regularization method for total variation-based image restoration*, Multiscale Model. Simul., 4 (2005), pp. 460–489.
- [38] X. OTAZU, M. GONZÁLEZ-AUDÍCANA, O. FORS, AND J. NÚÑEZ, *Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods*, IEEE Trans. Geosci. Remote Sensing, 43 (2005), pp. 2376–2385.
- [39] D. W. PEACEMAN AND H. H. RACHFORD, JR., *The numerical solution of parabolic and elliptic differential equations*, J. Soc. Indust. Appl. Math., 3 (1955), pp. 28–41.
- [40] E. RESMERITA AND O. SCHERZER, *Error estimates for non-quadratic regularization and the relation to enhancing*, Inverse Problems, 22 (2006), pp. 801–814.
- [41] S. SETZER, *Split Bregman algorithm, Douglas-Rachford splitting and frame shrinkage*, in SSVM '09: Proceedings of the Second International Conference on Scale Space and Variational Methods in Computer Vision, Berlin, Heidelberg, 2009, Springer-Verlag, Berlin, pp. 464–476.
- [42] V. SHAH., N. YOUNAN, AND R. KING, *An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets*, IEEE Trans. Geosci. Remote Sensing, 46 (2008), pp. 1323–1335.
- [43] V. K. SHETTIGARA, *A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set*, Photogrammetric Engrg. Remote Sensing, 58 (1992), pp. 561–567.
- [44] C. THOMAS, T. RANCHIN, L. WALD, AND J. CHANUSSET, *Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics*, IEEE Trans. Geosci. Remote Sensing, 46 (2008), pp. 1301–1312.
- [45] T.-M. TU, P. S. HUANG, C.-L. HUNG, AND C.-P. CHANG, *A fast intensity-hue-saturation fusion technique with spectral adjustment for IKONOS imagery*, IEEE Geosci. Remote Sensing Lett., 1 (2004), pp. 309–312.
- [46] U.S. ARMY CORPS OF ENGINEERS, *Urban dataset*, available online at <http://www.tec.army.mil/hypercube>.
- [47] V. VIJAYARAJ, C. G. O. HARA, AND N. H. YOUNAN, *Quality analysis of pansharpened images*, in Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IEEE, Washington, DC, 2004.

- [48] L. WALD, T. RANCHIN, AND M. MANGOLINI, *Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images*, Photogrammetric Engrg. Remote Sensing, 63 (1997), pp. 691–699.
- [49] Z. WANG AND A. C. BOVIK, *A universal image quality index*, IEEE Signal Process. Lett., 9 (2002), pp. 81–84.
- [50] R. H. YUHAS, A. F. H. GOETZ, AND J. W. BOARDMAN, *Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm*, in Proceeding Summaries of the 3rd Annual JPL Airborne Geoscience Workshop, 1992, pp. 147–149.
- [51] J. ZHOU, D. L. CIVICO, AND J. A. SILANDER, *A wavelet transform method to merge landsat TM and SPOT panchromatic data*, Internat. J. Remote Sensing, 19 (1998), pp. 743–757.