Variational Wavelet Pan-Sharpening

Michael Moeller, Todd Wittman, Andrea L. Bertozzi

***Abstract*—Pan-sharpening is the process of fusing a low resolution multispectral image with a high resolution panchromatic image to obtain a high resolution multispectral image. We propose a new pan-sharpening method called Variational Wavelet Pan-sharpening (VWP) that combines wavelet fusion and the edges of the panchromatic image as an energy minimization problem. Furthermore, we introduce additional energy terms to explicitly preserve the color information within each band and the correlation between bands. Numerical results are presented on Quickbird data and compared to various pan-sharpening methods with the help of image quality metrics. VWP is shown to outperform existing methods in the sense that it produces high quality images, whereas the choice of parameter determines the amount of spatial and spectral quality.**

***Index Terms*—image fusion, multispectral, pan-sharpening, variational, wavelet**

I. Introduction

MANY satellite imaging systems, including the Quickbird and Landsat-7 satellites, produce a panchromatic image to accompany the multispectral imagery. This panchromatic image has higher spatial resolution than the multispectral data, but the spectral response spans a wide range. For example, the Quickbird satellite produces 4-band multispectral images with 2.4 m resolution and panchromatic images with 0.6 m resolution. The panchromatic image can be used to enhance the resolution of the multispectral image through a process called sensor fusion or pan-sharpening.

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. This becomes especially important as the number of bands increases, because the spectral signature can be used for material identification. Therefore, the pan-sharpened image should possess both high spatial quality and spectral quality.

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) [1], [2], [3] and Brovey methods [4]. Other methods project the images into a different space like Principal Component Analysis (PCA) [5]. Several authors have proposed using the wavelet transform to extract geometric edge information from the panchromatic image [6], [7], [8]. Recently, Ballester et. al. proposed 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 [9].

There have been several survey papers comparing the performance of pan-sharpening methods [10], [11], [4]. The general conclusion is that there is a trade-off between spatial and spectral quality. For example, the classical IHS method produces images with excellent visual quality, but the pan-sharpened image contains noticeable spectral distortions. Our goal is to combine ideas from the various methods into a single variational framework to produce images that have both high spatial and spectral quality. We refer to our model as Variational Wavelet Pan-sharpening (VWP).

In Section II, we present an energy model that describes the desired qualities of the pan-sharpened image. We also present an alternate energy (AVWP) that allows for faster computation and produces similar results. We discuss numerical methods for minimizing the VWP energy in Section III. In Section IV, we present results on Quickbird data and compare VWP to existing pan-sharpening methods using a variety of image quality metrics. Finally, we conclude in Section V by suggesting extensions of VWP and areas for future research.

II. 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. One of the first and most famous variational methods is the Rudin-Osher-Fatemi Total Variation (TV) model [12]. A more general description of variational image processing methods can be found in [13].

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 generally true ([14], Fig. 1). A false linear combination assumption can lead to spectral distortion and therefore damage the spectral information of the original multispectral image. That is the reason why we propose a variational method that does not depend on this assumption. Another advantage of avoiding such a linear combination term is that our method can be extended to an arbitrary number of bands and is not restricted to those whose spectral response is covered by the panchromatic image.

The VWP energy functional consists of four parts and is specifically designed to sharpen images while preserving spectral quality. In the following we describe each term of our

Manuscript received December ?, 2008?

This work was supported by the US Department of Defense, ONR grant N000140810363, NSF grant ACI-0321917 and NSF grant DMS-0601395.

Michael Moeller is a student at the Westfaelische Wilhelms Universitaet Muenster, Germany, e-mail: michaelm@math.ucla.edu

Todd Wittman is with the Department of Mathematics, University of California, Los Angeles, CA, USA e-mail: wittman@math.ucla.edu

Andrea Bertozzi is with the Department of Mathematics, University of California, Los Angeles, CA, USA e-mail: bertozzi@math.ucla.edu

energy functional separately. We will refer to the panchromatic image as $P : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^2$ is the image domain. M_i are the low resolution multispectral image and u_i are the desired high resolution multispectral bands.

A. Geometry matching term

The first variational pan-sharpening method was proposed by Ballester et. al. in 2006 [9]. Their idea for introducing the geometry of the panchromatic image was to align all level lines of the high resolution panchromatic and each multispectral band. The main assumption 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. This vector field can be calculated almost everywhere under certain general assumptions. In practice the vector field is implemented as $\theta(x) = \frac{\nabla P(x)}{|\nabla P(x)|_\epsilon}$ where $|\nabla P|_\epsilon = \sqrt{(D_x P)^2 + (D_y P)^2 + \epsilon^2}$ is a regularization to avoid division by zero. The panchromatic image P then satisfies $\theta \cdot \nabla P = |\nabla P|$. To ensure that each multispectral band has the same level sets as the panchromatic image, they align the normal vectors of the level sets. Therefore every band of the restored image should satisfy $|\nabla u_n| - \theta \cdot \nabla u_n = 0$. The integral over the sum of these terms is added to the energy functional and integration by parts is applied to the second part of the integral. The total energy functional minimized by the P+XS model is

$$\begin{aligned} E(u_n) &= \sum_{n=1}^{4} \gamma_n \int_\Omega (|\nabla u_n| + div(\theta) \cdot u_n) dx \\ &\quad + \lambda \int_\omega (\sum_{n=1}^{4} \alpha_n u_n - P)^2 dx \\ &\quad + \mu \sum_{n=1}^{4} \int_\Omega \Pi_S((k_n * u_n) - M_n)^2 dx, \end{aligned} \quad (1)$$

where k_n is a convolution kernel, Π_S a Dirac comb and γ_n, λ, α_n and μ parameter to weight the different terms.

Unlike the original P+XS model, we want to be able to weight both parts of the geometry matching term separately by introducing parameter γ and η. This leads to the energy term

$$E_g = \sum_{n=1}^{N} [\gamma \int_\Omega |\nabla u_n| \, dx + \eta \int_\Omega div(\theta) \cdot u_n \, dx]. \quad (2)$$

This term does an excellent job enforcing spatial quality. Notice that the first part of this term is TV regularization [12] for each band. Furthermore, the idea of aligning the gradient vectors of an image with a smooth unit normal vector field was also proposed by [15] and lead to iterative regularization using the Bregman distance [16].

B. Wavelet matching term

The combination of wavelets and variational methods has recently been applied to many image processing tasks [17], [18], [19]. To match the colors of the low resolution multispectral image with sharper edges, we do a two level wavelet decomposition of the panchromatic image and each multispectral band. Then the high level wavelet coefficients are matched to the corresponding coefficients of the panchromatic image, while the low level approximation coefficient are matched to the low resolution multispectral band. Figure 1 illustrates the choice of the matching wavelet coefficients.

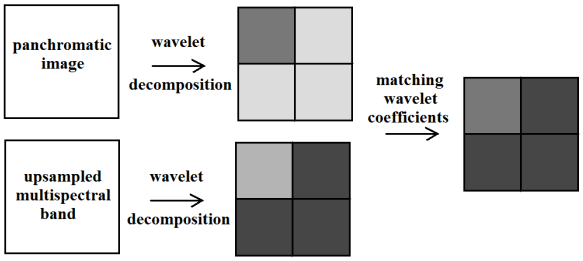

Fig. 1. Matching wavelet coefficients for a discrete wavelet transform

This choice of wavelet coefficients is well known in the literature and the same rule is used for wavelet pan-sharpening, e.g. in [6]. Many author have pointed out that the stationary (redundant) wavelet transform gives superior fusion results in comparison to the discrete wavelet transform ([20], [14], [21], [22]. The drawback of using stationary wavelets is the slower speed of the transform and the large increase in data.

To formalize this idea in a mathematical context for our energy functional we use the following 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 $1 \le k \le 3$, $j \in Z$ and $n = (n_1, n_2) \in Z^2$

$$\psi^k_{j,n}(x) = \frac{1}{2^j} \psi^k(\frac{x_1 - 2^j n_1}{2^j}, \frac{x_2 - 2^j n_2}{2^j}). \quad (3)$$

Further we define a two dimensional scaling function by

$$\phi^2_{j,n}(x) = \frac{1}{2^j} \phi(\frac{x_1 - 2^j n_1}{2^j}) \phi(\frac{x_2 - 2^j n_2}{2^j}). \quad (4)$$

Then the approximation coefficients of a two dimensional function are give by the scalar product with ϕ^2 and the 3 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

$$a^i_j[n] = <\uparrow M_i, \phi^2_{j,n}>, \quad (5)$$

where $\uparrow M_i$ denotes upsampling of the low resolution multispectral 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:

$$d_{\{k,j\}}[n] = < P, \psi^k_{j,n} >, \ for \ 1 \le k \le 3. \quad (6)$$

If we denote the desired approximation coefficients for band i by $\alpha^i_j[n]$ and the desired detail coefficients by $\beta^i_{\{k,j\}}[n]$ then

we add the following term to our energy functional:

$$E_w = \sum_n c_0(a_L^i[n] - \alpha_L^i[n])^2\phi_{j,n}^2 + \sum_n \sum_{j=1}^{L}\sum_{k=1}^{3} c_j(d_{\{k,j\}}[n] - \beta_{\{k,j\}}^i[n])^2\psi_{j,n}^k, \quad (7)$$

where c_0 is the parameter for the approximation coefficient matching, $c_j,\ 1 \le j \le L$ are the parameter for the different level of detail coefficient matching and L is the level of decomposition. In our experiments we used L=2. Notice that we assumed our continuous representations of the images are elements of $V_0^2 = span(\phi_{0,n}^2)$ since this holds in practice for the discrete formulation anyways. For $c_0 = c_1 = c_2$ this term would become a least squares match to a wavelet fused image. In our variational context we choose these parameter according to the type of image we would like to produce. For high spatial quality we especially want to introduce the edge information of the panchromatic image and therefore increase c_1 and c_2. Vice-versa we would choose larger values for c_0 for higher spectral quality.

C. Color preserving term

Using just the first two terms (2) and (8) gives very good spatial results, but we also want to enforce spectral quality. To preserve the spectral information within each band we would like to preserve the colors of the resized multispectral image at those parts of the image that have no edges or texture. We add the following term to our energy functional:

$$E_c = \nu \sum_{i=1}^{N} \int_{\Omega-\Gamma} (u_i - \uparrow M_i)^2\, dx. \quad (8)$$

Γ denotes the set of edges and texture in the panchromatic image which can be determined by any appropriate edge detector. In our experiments we calculated $\Gamma = exp(-\frac{d}{|\nabla P|^2})$ with a suitable constant d. This edge detector has been used successfully in many image processing applications such as the Perona-Malik model [23].

D. Spectral correlation preserving term

So far we constrain the colors within each band, but none of the terms couples the different bands. As mentioned earlier, a single pixel's spectral signature can be used for material classification, if the number of bands is high enough. In this case it is crucial to preserve the frequency information from the original low resolution multispectral image. To achieve this, we propose that every possible ratio of two different spectral bands of our pan-sharpened high resolution image should equal the ratio of the same bands of the original multispectral image. We would like to obtain at every pixel $\frac{u_i}{u_j} = \frac{\uparrow M_i}{\uparrow M_j} \Rightarrow u_i * \uparrow M_j - u_j * \uparrow M_i = 0$, where $\uparrow$ denotes upsampling to the high resolution image size. Therefore, we add the sum of the squares of the corresponding L_2 norms to our energy functional:

$$E_s = \mu \sum_{i,j=1 i<j}^{N} \int_{\Omega} (u_i \cdot \uparrow M_j - u_j \cdot \uparrow M_i)^2\, dx. \quad (9)$$

Another way of looking at this term is that it minimizes the spectral angle between each pixel frequency vector in the low resolution and in the sharpened multispectral image. For an arbitrary but fixed pixel let $\vec{a}$ be the frequency vector in the low resolution and $\vec{b}$ the frequency vector in the corresponding high resolution image. The minimizer of the above energy term is $a(i)\cdot b(j) - a(j)\cdot b(i) = 0\ \forall i,j$. This can be rewritten as $a(i) = \frac{a(j)}{b(j)}\cdot b(i) = 0\ \forall i,j$ which proves that $\vec{a}||\vec{b}$. This implies that the spectral angle $arccos(\frac{<\vec{a},\vec{b}>}{||\vec{a}||\cdot||\vec{b}||})$ is zero. The spectral angle is widely used to compare spectral information. Besides the well known spectral quality metric SAM [24] the spectral angle is also used in hyperspectral imaging for material comparison and classification [25].

The fitting terms described in C and D are new whereas E_g is a simple modification of the P+XS method and E_w puts the ideas of wavelet fusion similar to [6] with the ideas of [20] in a variational setting. The total energy functional can then be written as

$$E(u) = E_w + E_g + E_c + E_s. \quad (10)$$

This energy functional contains two different types of terms: three terms in the spatial domain and one matching term in the wavelet domain. Any minimization method will have to alternate between the wavelet and the spatial domain each iteration. This will slow down the whole algorithm significantly, especially for stationary wavelets. This is the reason why we propose an alternate energy, which can be minimized entirely in the spatial domain.

E. The Alternate Energy

The variational wavelet pan-sharpening method on the alternate energy (AVWP) is based on two ideas. First, if we choose our wavelet matching coefficients c_k equal for each level k, then the whole term is just a matching to a wavelet fused image. Second, away from the edges the matching to the low resolution multispectral image gives us the best color values we can get for our image. Therefore, we combine terms B and C to one matching term that matches the low resolution image away from edges and the wavelet fused image on the edges. Unlike variational segmentation algorithms like Mumford-Shah [26] the edgeset does not have to be evolved, since it is give by the panchromatic image. Again, the edge detection method we used for our algorithm is $exp(-\frac{d}{|\nabla P|^2})$.

Denoting the wavelet fused image for the q^{th} band with W_q and the new matching image with Z_q we have

$$Z_q = exp(-\frac{d}{|\nabla P|^2}) \cdot W_q + (1 - exp(-\frac{d}{|\nabla P|^2}))\cdot \uparrow M_q. \quad (11)$$

The terms I_c and I_w in our original energy are then replaced by

$$E_a = \nu \sum_{n=1}^{N} \int_{\Omega} (u_n - Z_n)^2\, dx. \quad (12)$$

AVWP minimizes the energy:

$$E(u) = E_g + E_s + E_a. \quad (13)$$

This energy can be minimized in the spatial domain only and will therefore allow for much faster computation.

III. Numerical Methods

For each of the two energy functionals (10) and (13) we implemented an explicit forward Euler gradient descend method and an alternating directions implicit (ADI) method [27], [28]. For the sake of simplicity, we only discuss the numerical details for the methods on our original energy functional. The implementations for the alternate energy are very similar.

One can show that minimizing the above energy functional (10) is equivalent to solving a partial differential equation (PDE), called the Euler-Langrange equation. We therefore calculate the first variations of our energy term. For the q^{th} band we get

$$\begin{aligned}
\frac{\delta E_g}{\delta u_q} &= -(\gamma div(\frac{\nabla u_q}{|\nabla u_q|_\epsilon}) - \eta div(\theta)) && (14)\\
\frac{\delta E_s}{\delta u_q} &= 2\mu \sum_{j=1, j\neq q}^{N} (u_q \cdot \uparrow M_j - u_j \cdot \uparrow M_q) \uparrow M_j && (15)\\
\frac{\delta E_c}{\delta u_q} &= 2\nu \chi_{\Omega-\Gamma}(u_q - \uparrow M_q) && (16)\\
\frac{\delta E_w}{\delta u_q} &= -2 \sum_n c_0 (a_L^q[n] - \alpha_L^q[n]) \phi_{j,n}^2 && (17)\\
&- 2 \sum_n \sum_j^L \sum_{k=1}^3 c_j (d_{\{k,j\}}[n] - \beta_{\{k,j\}}^q[n]) \psi_{j,n}^k,
\end{aligned}$$

where $\chi_{\Omega-\Gamma}$ denotes the characteristic function of $\Omega - \Gamma$ or a similar smooth edgeset indicator like $exp(-\frac{d}{|\nabla P|^2})$. The first variation of the energy functional must be zero in the minimum of the functional:

$$\frac{\delta E}{\delta u_q} = 0. \quad (18)$$

The corresponding minimizing argument is the pan-sharpened image and the solution we are looking for. To compute this solution we introduce an artificial time variable t, replace the zero on the right hand side by $-\frac{d}{dt}u_q$ and solve the resulting equation to steady state. This is called gradient descent or the method of steepest descent. In the following we present two different numerical schemes to perform the gradient descent.

A. Explicit Euler method

The time derivative $\frac{d}{dt}u_q$ is discretized explicitly. We then iteratively compute the solution by

$$\begin{aligned}
u_q^{t+1} &= u_q^t + \delta t * [(\gamma div(\frac{\nabla u_q^t}{|\nabla u_q^t|_\epsilon}) - \eta div(\theta))\\
&- 2\mu \sum_{j=1, j\neq q}^{N} (u_q^t \cdot \uparrow M_j - u_j^t \cdot \uparrow M_q) \uparrow M_j\\
&- 2\nu \chi_{\Omega-\Gamma}(u_q^t - \uparrow M_q)\\
&+ 2c_0 (a_L^q[n] - \alpha_L^{q,\,t}[n]) \phi_{j,n}^2\\
&+ 2 \sum_j^L \sum_{k=1}^3 c_j (d_{\{k,j\}}[n] - \beta_{\{k,j\}}^{q,\,t}[n]) \psi_{j,n}^k]. \quad (19)
\end{aligned}$$

We consider the iteration to have reached steady state when the relative change in energy between two iterates is less than 0.05%:

$$\frac{|E(u^{t+1}) - E(u^t)|}{E(u^t)} \cdot 100 \leq 0.05. \quad (20)$$

For the wavelet matching term we compute a wavelet decomposition of our current iteration, calculate $2\delta t[c_0(a_L^q[n] - \alpha_L^{q,\,t}[n]) + \sum_j^L \sum_{k=1}^3 c_j(d_{\{k,j\}}[n] - \beta_{\{k,j\}}^{q,\,t}[n])]$ and transform the result back to the spatial domain.

B. ADI method

For faster convergence of the numerical method we choose a different time discretization. First of all we split the wavelet matching term from the spatial terms by doing an alternating minimization. The minimization of the wavelet part is easy since we can directly compute the solution of the problem and just take one timestep towards it:

$$\begin{aligned}
\alpha_L^{q,\,t+1}[n] &= \alpha_L^{q,\,t}[n]\\
&+ 2 \cdot \delta t \cdot c_0 (a_L^q[n] - \alpha_L^{q,\,t}[n]), && (21)\\
\beta_{\{k,j\}}^{q,\,t+1}[n] &= \beta_{\{k,j\}}^{q,\,t}[n]\\
&+ 2 \cdot \delta t \cdot c_j (d_{\{k,j\}}[n] - \beta_{\{k,j\}}^{q,\,t}[n]). && (22)
\end{aligned}$$

For the minimization of the remaining spatial part, we apply an alternating directions minimization (ADI) method. This method is semi-implicit and requires two steps. In the first step we have to solve a linear equation for each row in the second for each column of the image. A more detailed description of the ADI scheme can be found in the appendix.

The numerical costs for calculating one timestep of the ADI scheme are much higher than for the explicit method, but the timestep for the ADI method can be chosen larger and the convergence properties are much better. In the next section, we show that the ADI method is much faster than the explicit timestepping scheme for both the original and alternate energies.

IV. Numerical Results

We implemented the ADI and implicit method as described above for AVWP and VWP in Matlab. Matlab's 'sym4' stationary wavelet transform was used in a second level wavelet decomposition.

We ran all methods on an Intel Duo Core processor with 2GHz and 3GB memory. The iteration is considered to have converged when equation 20 holds, i.e. the change in energy between two iterates is less than 0.05%. To compare the advantage in runtime of the ADI and the explicit method we took a square subsection of a Quickbird image with side length x for x $\in$ {0, 100, 200, ..., 600 } and ran the two minimization schemes. Figure (2) shows the results.

The ADI method is much faster for AVWP as well as for VWP. For large images, ADI on the original energy is even faster than the explicit scheme for the AVWP.

A great advantage of VWP and variational methods in general is their flexibility. PCA and IHS generally require no parameters and the only choice in wavelet fusion would be the

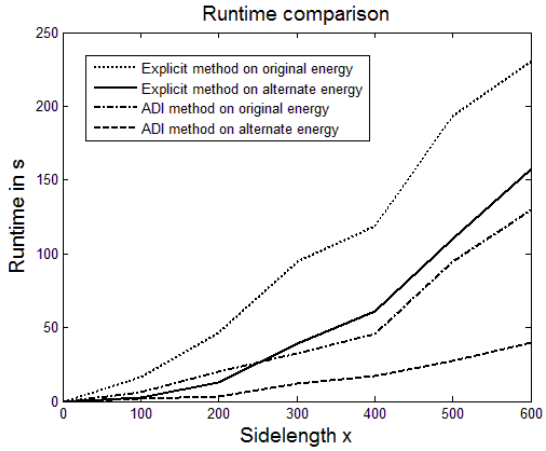

Fig. 2. Comparison of runtime

level of decomposition. In a variational method, the parameters have a huge influence on the result and thereby one can choose which term in the energy functional is more important for a certain purpose. Putting more weight on spectral terms will especially enforce spectral quality. On the other hand, using exactly the same model one can also focus on spatial quality and get results of spatial quality similar to IHS. Figure IV shows three different fusion results of the VWP method for different choices of parameter. This gives images ranging from very high spectral but low spatial quality to images with very high spatial but low spectral quality. Generally, there is a trade-off between spectral and spatial quality, but VWP has the flexibility to give almost any type of desired combination of spectral and spatial quality by adjusting the parameters.

Our standard choice of parameter for images with high spectral quality is given in table I. We will refer to this as

TABLE I
SPECTRAL QUALITY METRICS FOR IMAGE IV

	γ_n	ν	μ	ϵ	η	c_0	c_1	c_2
VWP	0.5	5	100	$10^{(-6)}$	0.5	4	2	2
AVWP	0.5	5	100	$10^{(-6)}$	0.5	-	-	-

parameter choice (1). For images with high spatial quality we choose the parameters shown in table II. These values will be

TABLE II
SPECTRAL QUALITY METRICS FOR IMAGE IV

	γ_n	ν	μ	ϵ	η	c_0	c_1	c_2
VWP	0.7	4	100	$10^{(-3)}$	1.4	0.5	4	4
AVWP	0.7	4	100	$10^{(-3)}$	1.4	-	-	-

referred to as parameter choice (2).

Figure IV shows the result of the classical IHS, the stationary wavelet transform and the VWP method for these two different choices of parameter.

One can see that the spatial quality of the IHS fused images is very high. All objects have very sharp edges and the whole image seems to be very close to the panchromatic image in terms of spatial quality. However, one can also see that the colors of the IHS fused images slightly changed in comparison to the original image (e.g. the color of the trees in figure IV). Color distortion means that the frequency information stored in the multispectral image is lost or at least greatly damaged. The VWP image with high spatial quality seems to be closer to the original multispectral image in terms of colors and also has spatial quality comparable if not better than IHS. The other VWP result clearly looks blurry in comparison and is similar to the wavelet fused image. Unfortunately, spectral quality is much harder to see than spatial quality. Large spectral distortion like in the IHS image can be seen in the change of color, but when the spectral distortion is a little smaller, it is almost undetectable visually. Therefore, we take several quality metrics into account to help us with the evaluation.

In [29] the following seven spectral quality metrics were implemented to judge the performance of pan-sharpening methods:

- The idea of the **relative dimensionless global error in synthesis (ERGAS)** is to take the average mean square error normalized by the mean of each band [30].
- The **Spectral Angle Mapper (SAM)** calculates the average change in angle of all frequency vectors [24].
- The **Spectral Information Divergence (SID)** views each pixel spectrum as a random variable and then measures the discrepancy of probabilistic behaviors between spectra [31].
- The **Universal Image Quality Index (Q-average)** models distortion as a combination of three different factors: loss of correlation, luminance distortion, and contrast distortion. The best value for Q-average is 1 [32].
- The **Relative Average Spectral Error (RASE)** calculates the mean error of all bands per radiance of the image [33].
- **Mean Change in Correlation Coefficients (MCCC)** is the average of the absolute values of the change in correlation coefficients between band i and band j $\forall i, j$ before and after the sharpening process [34].
- **Root mean squared error (RMSE)** is the average squared difference between the original multispectral and pan-sharpened image.

The results of the evaluation of all quality metrics for Intensity-Hue-Saturation Fusion (IHS), Principal Component Analysis (PCA), discrete wavelet fusion (DWave), stationary wavelet fusion (SWave), P+XS image fusion, VWP and AVWP for our testimage IV are shown in table III.

To get more meaningful and less image dependent quality metric values we ran all methods on five different images and evaluated each fusion result with the quality metrics. The average of all five images of all method is shown in table IV. The best metric quality results are shown in bold print.

To a large extent the quality metrics confirm our visual analysis. Clearly there is a trade-off between spatial and spectral quality. The VWP fusion result where we focused on spatial quality (AVWP$^+$) has low quality metric values but therefore provides visual quality equal or better than the IHS fused image. If we look at fig. IV the colors of the AVWP$^+$ seem to be closer to the original low resolution image than the IHS image. In the IHS image the colors of the trees appear in a lighter green than the low resolution image shows. VWP seems to preserve these colors much better.

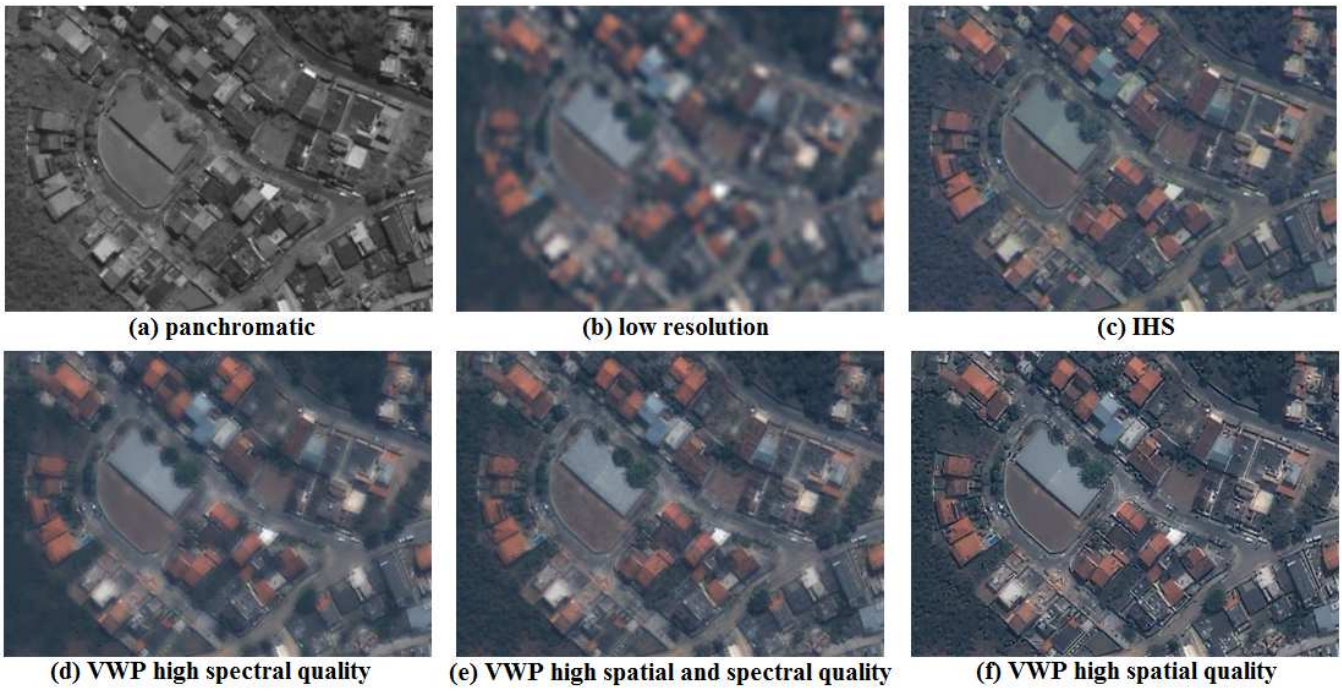

Fig. 3. Effect of the choice of parameter in VWP; (a), (b): Copyright DigitalGlobe, NextView liscense, 2003; image (d): parameter choice (1), image (f): parameter choice (2), image (e): compromise between (d) and (f) for spatial and spectral quality

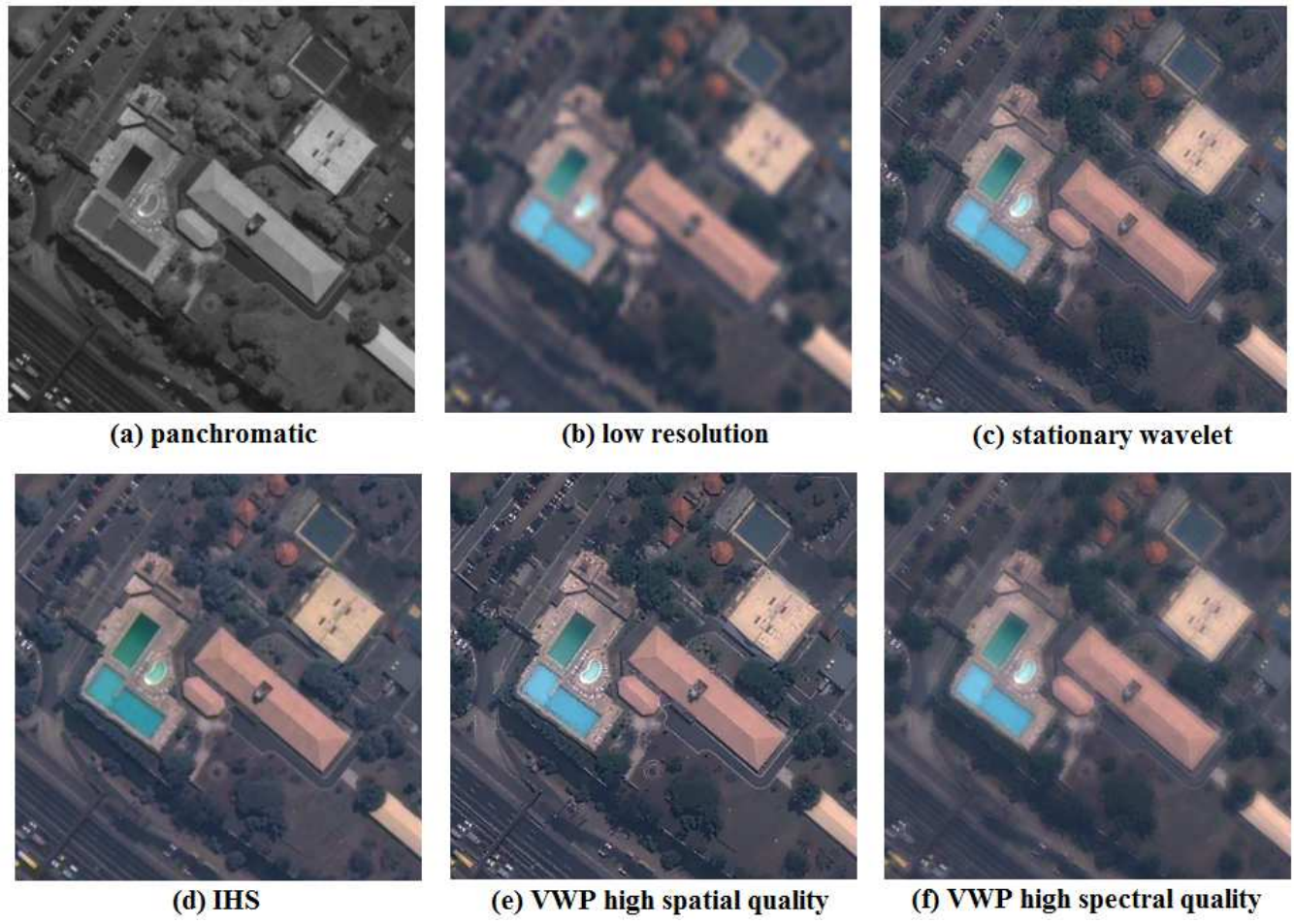

Fig. 4. Results of different pan-sharpening methods; (a), (b): Copyright DigitalGlobe, NextView liscense, 2003; image (e): parameter choice (2), image (f): parameter choice (1)

TABLE III
SPECTRAL QUALITY METRICS FOR IMAGE IV

	$MCCC^*$	ERGAS	Qave	RASE	RMSE	SAM	SID^{**}
IHS	6.68	3.41	0.992	13.0	53.3	2.02	21.3
PCA	16.5	4.96	0.951	19.6	80.2	4.40	18.3
D. Wav.	3.04	2.49	0.977	10.00	41.0	2.68	17.3
S. Wav.	1.67	1.92	0.991	7.75	31.8	1.39	1.82
P+XS	1.94	2.46	0.986	9.45	38.7	1.54	3.42
VWP	**0.998**	**1.40**	**0.997**	**5.63**	**23.1**	0.366	0.401
AVWP	1.12	1.44	0.997	5.82	23.9	**0.325**	**0.336**
AVWP$^+$	6.10	4.342	0.981	13.84	56.74	0.465	4.13

*: given in $10^{(-2)}$ **: given in $10^{(-3)}$ $^+$: good spatial quality

TABLE IV
AVERAGE SPECTRAL QUALITY METRICS FOR 5 IMAGES

	$MCCC^*$	ERGAS	Qave	RASE	RMSE	SAM	SID^{**}
IHS	5.19	3.31	0.994	12.6	39.3	1.56	7.51
PCA	9.37	3.95	0.968	15.4	49.8	2.99	13.0
D. Wav.	3.73	2.43	0.983	9.66	30.5	2.06	8.87
S. Wav.	2.37	1.99	0.993	7.88	24.5	1.10	1.97
P+XS	2.54	2.92	0.983	10.6	29.9	1.96	5.51
VWP	**1.40**	**1.45**	**0.997**	**5.82**	**18.1**	0.464	0.737
AVWP	1.80	1.55	**0.997**	6.22	19.4	**0.338**	**0.315**
AVWP$^+$	6.80	3.26	0.985	13.10	42.5	0.488	12.2

*: given in $10^{(-2)}$ **: given in $10^{(-3)}$ $^+$: good spatial quality

The stationary wavelet fusion and P+XS give good quality metric values, but their visual spatial quality is not as good as AVWP$^+$. Their fusion results look more blurry.

The different VWP images in fig. IV show the range of results we can produce. While one of them is providing spatial quality equal if not better than the IHS method and seemed to have the colors preserved much better, the other VWP result shows how much spectral quality we can preserve. The middle image is a compromise between spatial and spectral quality: It has sharper edges and more texture than the high spectral quality VWP and more spectral information than the high spatial quality VWP. For the high spectral quality VWP, we obtain the best quality metric values for all 7 spectral quality metrics. Especially the SAM quality metric which we directly wanted to minimize in our energy approach is more than 3 times better in AVWP than in any other method. Since SAM is (as mentioned earlier) also used for material classification in higher dimensional imagery this choice of parameter seems to be very well suited for all applications, where the spectrum of each pixel is important. Even in the the high spatial quality AVWP$^+$ image the SAM quality metric is still more than factor of 2 times better than stationary wavelet fusion which is the second best method in this category.

In summary we can say that VWP can produce a very wide range of good fusion results. VWP can produce images with spatial quality equal to IHS while still having more realistic colors or spectral qualities better than any other fusion method. There is definitely a trade-off between spectral and spatial quality, but VWP has the flexibility to give almost any type of desired combination of spectral and spatial quality by adjusting the parameters.

V. CONCLUSION

We proposed a variational method based on the ideas of P+XS and wavelet image fusion for the task of pan-sharpening multispectral images. The model incorporates the alignment of all unit normal vectors of the level sets of each band with the panchromatic image. It includes color and edge matching in the wavelet domain and further preserves frequency information by keeping the ratio of all bands constant and by additionally matching the colors away from edges and texture.

VWP can produce a wide range of images where the user can decide about the importance of spectral and spatial quality by adjusting the parameter. High spatial and high spectral quality images were shown. We implemented eight different quality metrics to evaluate the performance of our proposed method in comparison to the most common other pan-sharpening methods. Our method seems to be the best choice if one wants to preserve the spectral information from the multispectral image.

Another advantage of our method in comparison to other methods is that it can be extended to an arbitrary number of spectral bands. If a panchromatic image was available we could sharpen hyperspectral images with 100-200 bands. VWP is particularly well-suited for this situation because it explicitly preserves spectral information.

For future research one could look into other numerical schemes for faster minimization, such as graph cuts or operator splitting methods. Technical information about the satellite sensors could be incorporated into the variational framework. Linked to this would be the question about the fusion of other totally different types of sensor data such as SAR images or MRI data for medical imagery.

APPENDIX
ALTERNATING DIRECTIONS IMPLICIT METHOD

The remaining spatial part of the gradient descent after the splitting is the following:

$$\begin{aligned}\frac{d}{dt}u_q &= (\gamma div(\frac{\nabla u_q}{|\nabla u_q|_\epsilon}) - \eta div(\theta)) \\ &- 2\mu \sum_{j=1,j\neq q}^{N} (u_q \cdot \uparrow M_j - u_j \cdot \uparrow M_q) \uparrow M_j \\ &- 2\nu\chi_{\Omega-\Gamma}(u_q - \uparrow M_q) \\ &+ 2\sum_n c_0(a_L^q[n] - \alpha_L^q[n])\phi_{j,n}^2. \end{aligned} \quad (23)$$

We discretize the derivatives in the numerator of the TV term implicitly in the x-direction and explicitly in the y-direction for half of a timestep and vice-versa for another half of a timestep. We get a linear equation for each row of our image for the implicit x-discretization and for each column for the implicit y-discretization both with tridiagonal matrices. For the step implicit in x direction we get:

$$\begin{aligned}u_q^{t+1} &= u_q^t + \delta t * [D_x(C^t \cdot D_x u_q^{t+1}) + D_y(C^t \cdot D_y u_q^t) \\ &- 2\mu \sum_{j=1,j\neq q}^{N} (u_q^{t+1} \cdot \uparrow M_j - u_j^t \cdot \uparrow M_q) \uparrow M_j \\ &- 2\nu\chi_{\Omega-\Gamma}(u_q^{t+1} - \uparrow M_q) - \eta * div(\theta)]. \end{aligned} \quad (24)$$

With the notation $C^t = \frac{\gamma}{|\nabla u_q^t|_\epsilon}$. We put all terms that include u^{t+1} on the left-hand side and all others on the right. Using backward differences for the first D_x and forward differences for the second D_x the following equation holds at every pixel:

$$\begin{aligned}[D_x(C^t \cdot D_x u_q^{t+1})]_{i,j} &= C_{i,j}^t u_{q;\,i,j+1}^{t+1} \\ &- (C_{i,j}^t + C_{i,j-1}^t) u_{q;\,i,j}^{t+1} \\ &+ C_{i,j-1}^t u_{q;\,i,j-1}^{t+1}. \end{aligned} \quad (25)$$

We can now solve a linear equation for each row of our image. For an nxm image we define the tridiagonal matrix A^i by

$$A^i = \begin{pmatrix} b_1^i & c_1^i & 0 & 0 & ... & 0 \\ a_2^i & b_2^i & c_2^i & 0 & ... & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & ... & 0 & a_{m-1}^i & b_{m-1}^i & c_{m-1}^i \\ 0 & ... & 0 & 0 & a_m^i & b_m^i \end{pmatrix}, \quad (26)$$

with

$$\begin{aligned} a_1^i &= 0, \\ a_j^i &= -\delta t\, C(i, j-1), \\ c_m^i &= 0, \\ c_j^i &= -\delta t\, C(i, j+1), \\ b_j^i &= 1 + 2\delta t\nu\chi_{\Omega-\Gamma} + 2\mu\delta t \sum_{k=1,k\neq q}^{N} (M_k(i,j))^2 - c_j^i - a_j^i. \end{aligned} \quad (27)$$

It is easy to see that each matrix A^i is diagonal dominant, which allows us to solve each linear equation very efficiently with the Thomas (TDMA) algorithm [35]. The procedure for the step taken implicitly in the y-direction is similar. Here we solve a linear equation for each column of the image. For both steps we assumed zero Neumann boundary conditions for our image.

ACKNOWLEDGMENT

The authors would like to thank M. Strait, S. Rahmani and D. Merkurev for their help and especially for provinding the Matlab code for all image quality metrics. We would further like to thank J. Dobrosotskaya, Prof. S. Osher, Dr. J. Darbon, E. Esser and Prof. M. Burger for their advice.

REFERENCES

[1] M. Choi, H.-C. Kim, N. Cho, and H. O. Kim, "An Improved Intensity-Hue-Saturation Method for IKONOS Image Fusion," *submitted to IJRS*.
[2] M. Choi, "A New Intensity-Hue-Saturation Fusion Approach to Image Fusion with a Tradeoff Parameter," *IEEE Trans. of Geosci. and Remote Sens.*, vol. 44, pp. 1672–1682, 2006.
[3] 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. and Remote Sens. Lett.*, vol. 1, no. 4, October 2004.
[4] Q. Du, N. Younan, R. King, and V. Shah, "On the performance evaluation of pan-sharpening techniques," *IEEE Geosci. and Remote Sens. Lett.*, vol. 4, no. 4, pp. 518–522, October 2007.
[5] V. K. Shettigara, "A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set," *Photogrammetric eng. and remote sens.*, vol. 58, no. 5.
[6] J. Zhou, D. L. Civico, and J. A. Silander, "A wavelet transform method to merge Landsat TM and SPOT panchromatic data," *Int. J. of Remote Sens.*, vol. 19, no. 4, 1998.
[7] R. King and J. Wang, "A wavelet based algorithm for pan sharpening landsat 7 imagery," *Geosci. and Remote Sens. Symp., 2001. IGARSS '01. IEEE 2001 Int.*, vol. 2, pp. 849–851, 2001.
[8] N. Memarsadeghi, J. Le Moigne, and D. Mount, "Image fusion using cokriging," *Geosci. and Remote Sens. Symp., 2006. IGARSS 2006. IEEE Int. Conf. on*, pp. 2518–2521, August 2006.
[9] C. Ballester, V. Caselles, L. Igual, and J. Verdera, "A Variational Model for P+XS Image Fusion," *Int. J. of Comput. Vision*, vol. 69, no. 1, pp. 43–58, August 2006.
[10] V. Vijayaraj, C. G. O. Hara, and N. H. Younan, "Quality Analysis of Pansharpened Images," *Geosci. and Remote Sens. Symp., 2004. IGARSS '04. Proc. 2004 IEEE Int.*, vol. 1, September 2004.
[11] Q. Du, O. Gungor, and J. Shan, "Performance evaluation for pan-sharpening techniques," *Geosci. and Remote Sens. Symp., 2005. IGARSS '05. Proc. 2005 IEEE Int.*, vol. 6, pp. 4264–4266, July 2005.
[12] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear Total Variation Based Noise Removal Algorithms," *Physica D*, vol. 60, pp. 259–268, November 1992.
[13] T. F. Chan and J. Shen, *Image Processing and Analysis*. SIAM, 2005.
[14] X. Otazu, M. Gonzlez-Audcana, O. Fors, and J. Nez, "Introduction of Sensor Spectral Response Into Image Fusion Methods. Application to Wavelet-Based Methods," *IEEE Trans. on Geosci. and Remote Sens.*, vol. 43, no. 10, October.
[15] O. Lysaker, S. Osher, and X.-C. Tai, "Noise removal using smoothed normals and surface fitting," *IEEE Trans. on Image Process.*, vol. 13, no. 10, pp. 1345–1357, 2004.
[16] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iteratice regularization method for total variation-based image restoration," *SIAM Multiscale Modeling and Simulation*, vol. 4, pp. 460–489, 2005.
[17] T. F. Chan, J. Shen, and H.-M. Zhou, "A Total Variation Wavelet Inpainting Model with Multilevel Fitting Parameters," *Congr., Advanced signal process. algorithms, architectures, and implementations XVI*, vol. 69, no. 1, August 2006.
[18] J. Dobrosotskaya and A. Bertozzi, "A Wavelet-Laplace Variational Technique for Image Deconvolution and Inpainting," *IEEE Trans. on Image Process.*, vol. 17, pp. 657–663, May 2008.
[19] F. Malgouyres, "Mathematical Analysis of a Model Which Combines Total Variation and Wavelet for Image Restoration," *J. of inform. processes*, vol. 2:1, pp. 1–10, 2002.

[20] Y. Chibani and A. Houacine, "Redundant versus orthogonal wavelet decomposition for multisensor image fusion," *Pattern Recognition*, vol. 36, no. 4, pp. 879–887, April 2002.
[21] G. Hong and Y. Zhang, "The Effects of Different Types of Wavelets on Image Fusion," *Proc. Int. Conf. on Image Process.*, vol. 3, pp. 248 – 251, October 1995.
[22] 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 Geosci. and Remote Sens.*, vol. 40, pp. 2300 – 2312, 2002.
[23] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, July 1990.
[24] Goetz, Boardman, and Yunas, "Discrimination Among Semi-Arid Landscape Endmembers Using the Spectral Angle Mapper(SAM) Algorithm," *Proc. Summeries 3rd Annu. JPL Airborne Geosci. Workshop*, pp. 147 – 149, 1992.
[25] P. Shippert, "Introduction to Hyperspectral Image Analysis," *Online J. of Space Commun.*, 2003.
[26] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Commun. on Pure and Appl. Math.*, vol. 42, no. 5, pp. 577–685, 1989.
[27] D. W. Peaceman and J. H. H. Rachford, "The numerical solution of parabolic and elliptic differential equations," *J. of the Soc. for Ind. and Applied Math.*, vol. 3, no. 1, pp. 28–41, 1955. [Online]. Available: http://link.aip.org/link/?SMM/3/28/1
[28] D. W. P. Jim Douglas Jr., "Numerical solution of two-dimensional heat-flow problems," *AIChE Journal*, vol. 1, no. 4, pp. 505–512, 1955.
[29] M. Strait, S. Rahmani, and D. Merkurev, "Evaluation of Pan-Sharpening Methods," *UCLA Tech. Rep.*, August 2008.
[30] L. Wald, "Quality of high resolution synthesized images: Is there a simple criterion?" *Proc. of the 3rd conf. "'Fusion of Earth data: merging point measurements, raster maps and remotely sensed images"'*, January 2000.
[31] C.-I. Chang, "Spectral information divergence for hyperspectral image analysis," *Geosci. and Remote Sens. Symp., 1999. IGARSS '99 Proc. IEEE 1999 Int.*, vol. 1, pp. 509–511, 1999.
[32] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81 – 84, March 2002.
[33] M. Choi, "A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter," *IEEE Trans. on Geosci. and Remote Sens.*, vol. 44, no. 6, pp. 1672–1682, June 2006.
[34] V. Vijayaraj, C. G. O. Hara, and N. H. Younan, "Quality Analysis of Pansharpened Images," *Geosci. and Remote Sens. Symp., 2004. IGARSS '04. Proc. 2004 IEEE Int.*, vol. 1, September 2004.
[35] S. Conte and C. deBoor, *Elementary Numerical Analysis*. McGraw-Hill, 1972.

Michael Moeller is a student at the Westfaelische Wilhelms Universitaet Muenster, Germany. He is currently participating in a one year exchange program at the University of California, Los Angeles. His research interests lie in PDEs and the calculus of variations and their application to image processing.

Todd Wittman received the B.A. from Dartmouth College in 1997 and the Ph.D. in mathematics from University of Minnesota in 2006. He is currently a postdoctoral researcher in the mathematics department at University of California, Los Angeles. His research interests include PDEs and the calculus of variations and their application to problems in image processing.

Andrea L. Bertozzi received the B.A., M.A., and Ph.D. degrees in mathematics from Princeton University, Princeton, NJ, in 1987, 1988, and 1991 respectively. She was on the faculty of the University of Chicago, Chicago, IL, from 1991-1995 and Duke University, Durham, NC, from 1995-2004. During 1995-1996, she was the Maria Goeppert-Mayer Distinguished Scholar at Argonne National Laboratory. Since 2003, she has been with the University of California, Los Angeles, as a Professor of mathematics and currently serves as the Director of Applied Mathematics. Her research interests include image inpainting, image segmentation, cooperative control of robotic vehicles, swarming, and fluid interfaces. Prof. Bertozzi is a member of the Society for Industrial and Applied Mathematics, the American Mathematical Society, and the American Physical Society. She has served as a Plenary/Distinguished Lecturer for both SIAM and AMS and is an Associate Editor for the SIAM journals Multiscale Modelling and Simulation, Mathematical Analysis, and SIAM Review (Survey and Reviews). She also serves on the editorial board of Interfaces and Free Boundaries, Applied Mathematics Research Express, and Nonlinearity. Her past honors include a Sloan Foundation Research Fellowship and the Presidential Career Award for Scientists and Engineers.