Spectrally Consistent Satellite Image Fusion with Improved Image Priors

H. Aanæs, A. A. Nielsen and T. Bøvith
Informatics and Mathematical Modeling
Technical University of Denmark
DK-2800 Kgs. Lyngby, Denmark
{haa,aa,thb}@imm.dtu.dk

J. R. Sveinsson and J. A. Benediktsson
Department of Electrical and Computer Engineering
University of Iceland
Hjardarhaga 2-6, 107 Reykjavik
{sveinsso, benedikt}@hi.is

Abstract—Here an improvement to our previous framework for satellite image fusion is presented. A framework purely based on the sensor physics and on prior assumptions on the fused image. The contributions of this paper are two fold. Firstly, a method for ensuring 100% spectrally consistency is proposed, even when more sophisticated image priors are applied. Secondly, a better image prior is introduced, via data-dependent image smoothing.

I. Introduction

Image fusion is the subset of data fusion dealing with merging images. The most common meaning of "quality" in image fusion has been visual improvement, i.e., taking into account how the human perceives the fusion product. A frequently used meaning of "greater quality" is improved classification accuracy, in particular automated classification, an example is the correct labeling of crops and/or urban objects. Many other objectives have also been identified, see [4], [6].

Image fusion can be done at several levels: Pixel level, feature level, object level and decision level, depending on the intended use of the fused image. This paper is only concerned with pixel level fusion and when the terms "image fusion" or "fusion" are used, pixel level fusion is intended. In the current context fusion is the next step after preprocessing and the step before further processing, such as segmentation and classification. Other work in this area comprise [1], [2], [5], [7].

In [8] a new fusion algorithm, called SCP (Spectrally Consistent Pansharpening), which utilizes the known spectral responses of the photosensors to fuse multispectral and panchromatic satellite images at the pixel level into a high-resolution multispectral image, was proposed. The intention there was to create an image using as much information as feasible from all the available images and all available *a priori* knowledge. The image obtained by SCP maintain spectral consistency with the original multispectral image. Then a more elaborate model for the imaging physics was introduced and a fusion framework was developed in which it is possible to, simultaneously, keep spectral consistency and smooth the final image.

This paper is an improvement to SCP, in which a better image prior is introduced, via data-dependent image smoothing, and a method for maintaining spectral consistency using such sophisticated image priors.

II. The Sensor Fusion Problem

The problem under consideration is that of estimating a high resolution RGB-image, $(\mathbf{R}^{high}, \mathbf{G}^{high}, \mathbf{B}^{high})$, from a low resolution RGB-image , $(\mathbf{R}^{low}, \mathbf{G}^{low}, \mathbf{B}^{low})$, and a high resolution pan-chromatic image, $\mathbf{P}^{high}$. For ease of presentation we are going to focus on the IKONOS earth imaging satellite, where the ratio between high and low resolution is 4 by 4 to 1 i.e. 1:16. This is however done without loss of generality, as is also the case with the 'choice' of given images/channels /bands. We will also assume a perfect alignment of the high and low resolution images, such that a high resolution pixel is associated with one low resolution only.

A. Observation Model

As noted in our prior work [8], the value of a given pixel is determined as a spectral response function, $F(w)$[1] convolved with the light, $L(\mathbf{x}, \omega)$, hitting the associated patch, A, on the sensor, i.e.

$$\int_A \int_\Omega L(\mathbf{x}, \omega) F(w) d\omega d\mathbf{x} \ , \qquad (1)$$

where $\mathbf{x}$ denotes position and ω frequency, the latter running over a relevant spectral range Ω. This induces

---

[1]$F(\omega)$ is usually supplied by the sensor manufacture.

an inner product, e.g.[2]

$$< F^{red}, F^{pan} >= \int_{\Omega} F^{red}(\omega) F^{pan}(w) d\omega \ . \quad (2)$$

Thus a measure of how much information about e.g. $\mathbf{G}^{high}$ is contained in $\mathbf{P}^{high}$ is given by

$$< G, P >= \frac{< F^{green}, F^{pan} >}{\sqrt{< F^{green}, F^{green} >< F^{pan}, F^{pan} >}} \ , \quad (3)$$

which is equal to the cosine of the angle between the two spectra. This was the core of our work in [8], where an 'information' matrix was formed

$$\Sigma = \begin{bmatrix} < R, R > & < R, G > & < R, B > & < R, P > \\ < G, R > & < G, G > & < G, B > & < G, P > \\ < B, R > & < B, G > & < B, B > & < B, P > \\ < P, R > & < P, G > & < P, B > & < P, P > \end{bmatrix}$$

The problem was then cast in a statistical normal distribution setting with Σ as the variance and the low resolution RGB-values as means for the corresponding high resolution values. The estimate of $(\mathbf{R}^{high}, \mathbf{G}^{high}, \mathbf{B}^{high})$ were then found as the maximum likelihood estimate of their conditional distributions given $\mathbf{P}^{high}$. Specifically, the estimator is given by

$$\begin{bmatrix} \widehat{\mathbf{R}^{high}_{ij}} \\ \mathbf{G}^{high}_{ij} \\ \mathbf{B}^{high}_{ij} \end{bmatrix} = \mathbf{D}_{ij} = \begin{bmatrix} \mathbf{R}^{low}_i \\ \mathbf{G}^{low}_i \\ \mathbf{B}^{low}_i \end{bmatrix} + \begin{bmatrix} < R, P > \\ < G, P > \\ < B, P > \end{bmatrix} (\mathbf{P}^{high}_{ij} - P^{\mu}_i) \ , \quad (4)$$

where the high resolution pixels associated with the low resolution pixel $\mathbf{X}^{low}_i$ is denoted by $\mathbf{X}^{high}_{ij}$, $j \in \{1 \dots 16\}$. Here P^{μ}_i denotes the mean of the $\mathbf{P}^{high}_{ij}$ w.r.t. the low resolution pixel i. See Section VI for an example.

B. Spectral Consistency

Another implication of (1) is that of spectral consistency. Consider e.g. the blue channel and let the areas A_i and A_{ij} be associated with $\mathbf{B}^{low}_i$ and $\mathbf{B}^{high}_{ij}$, respectively. Then

$$\mathbf{B}^{low}_i = \int_{A_i} \int_{\Omega} L(\mathbf{x}, \omega) F(w) d\omega d\mathbf{x} =$$

$$\sum_{j=1}^{16} \int_{A_{ij}} \int_{\Omega} L(\mathbf{x}, \omega) F(w) d\omega d\mathbf{x} = \sum_{j=1}^{16} \mathbf{B}^{high}_{ij} \ .$$

So in general

$$\begin{bmatrix} \mathbf{R}^{low}_i \\ \mathbf{G}^{low}_i \\ \mathbf{B}^{low}_i \end{bmatrix} = \sum_{j=1}^{16} \begin{bmatrix} \mathbf{R}^{high}_{ij} \\ \mathbf{G}^{high}_{ij} \\ \mathbf{B}^{high}_{ij} \end{bmatrix} \ . \quad (5)$$

This spectral consistency is a very important property for image fusion algorithms applied to satellite images. This property is implicitly obtained by our previous method described above.

C. Regularization

The high resolution RGB-image, $(\mathbf{R}^{high}, \mathbf{G}^{high}, \mathbf{B}^{high})$, is not fully observable from the data, and hence regularization — explicit or implicit — is required to determine the remaining degrees for freedom. The estimate in (4) implicitly penalizes the estimate in the 2-norm, which gives a rather blocky effect corresponding to the low resolution blocks. In [8] a smoothness prior was also attempted, at the expense of spectral consistency, giving a somewhat blurred image if the blockyness should be addressed satisfactory. This smoothness prior has the form

$$\sum_{ij} \sum_{k \in \mathcal{N}_{ij}} \left\| \begin{bmatrix} \mathbf{R}^{high}_{ij} \\ \mathbf{G}^{high}_{ij} \\ \mathbf{B}^{high}_{ij} \end{bmatrix} - \begin{bmatrix} \mathbf{R}^{high}_k \\ \mathbf{G}^{high}_k \\ \mathbf{B}^{high}_k \end{bmatrix} \right\|_2^2 \ , \quad (6)$$

where $\mathcal{N}_{ij}$ denotes the neighboring pixels of pixel ij. Denoting the deviation by

$$\epsilon_{ijk} = \begin{bmatrix} \mathbf{R}^{high}_{ij} \\ \mathbf{G}^{high}_{ij} \\ \mathbf{B}^{high}_{ij} \end{bmatrix} - \begin{bmatrix} \mathbf{R}^{high}_k \\ \mathbf{G}^{high}_k \\ \mathbf{B}^{high}_k \end{bmatrix} \ ,$$

the smoothness term, (6), can be reformulated as

$$\sum_{ij} \sum_{k \in \mathcal{N}_{ij}} \rho_2(\epsilon_{ijk}) \ , \quad (7)$$

where ρ_2 denotes the error function, in this case the 2-norm. To use this in conjunction with (4), the latter has to be formulated as an error function, this was done as follows[3]

$$\sum_{ij} \mathcal{D}_{ij} = \quad (8)$$

$$\sum_{ij} \left(\begin{bmatrix} \mathbf{R}^{high}_{ij} \\ \mathbf{G}^{high}_{ij} \\ \mathbf{B}^{high}_{ij} \end{bmatrix} - \mathbf{D}_{ij} \right)^T \Sigma_C^{-1} \left(\begin{bmatrix} \mathbf{R}^{high}_{ij} \\ \mathbf{G}^{high}_{ij} \\ \mathbf{B}^{high}_{ij} \end{bmatrix} - \mathbf{D}_{ij} \right) \ ,$$

where Σ_C is the appropriate variance structure, derived from the conditional distribution of $\mathbf{D}_{ij}$. Thus the combined error function to be minimized is

$$\min_{\mathbf{R}^{high}, \mathbf{G}^{high}, \mathbf{B}^{high}} \sum_{ij} \left(\mathcal{D}_{ij} + \gamma \sum_{k \in \mathcal{N}_{ij}} \rho_2(\epsilon_{ijk}) \right) \ , \quad (9)$$

---

[2]The response functions are independent of position, so there is no need to integrate over area.

[3]The letter 'D' with different styles denotes two different entities, where $\mathcal{D}_{ij}$ is short hand for the data term.

where γ is a tuning parameter encapsulating the weight of the smoothness term relative to the data term. In [8] this was done via Markov random fields and simulated annealing.

III. Ensuring Spectral Consistency

The first problem we address in this paper, is how to maintain spectral consistency, i.e. (5), while smoothing the image or using other more sophisticated regularization methods or priors. We propose to do this by reparameterizing the problem such that only spectrally consistent solutions are possible. This is done by constraining the solution to the plane defined by (5), which for each low resolution pixel is a 15 dimensional space. For each low resolution pixel, i, define

$$\mathcal{R}_{il}^{high}, \; \mathcal{G}_{il}^{high}, \; \mathcal{B}_{il}^{high}, \quad l \in \{1 \ldots 15\} \; ,$$

which is our new representation of the problem. Define the linear mapping from our new parameter space, $\Re^{15}$, to our old, $\Re^{16}$, as

$$\begin{bmatrix} \mathbf{R}_{ij}^{high} \\ \mathbf{G}_{ij}^{high} \\ \mathbf{B}_{ij}^{high} \end{bmatrix} = \begin{bmatrix} \mathcal{R}_{il}^{high} \\ \mathcal{G}_{il}^{high} \\ \mathcal{B}_{il}^{high} \end{bmatrix} , \quad j \in \{1 \ldots 15\}$$

$$\begin{bmatrix} \mathbf{R}_{i,16}^{high} \\ \mathbf{G}_{i,16}^{high} \\ \mathbf{B}_{i,16}^{high} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_i^{low} \\ \mathbf{G}_i^{low} \\ \mathbf{B}_i^{low} \end{bmatrix} - \sum_{l=1}^{15} \begin{bmatrix} \mathcal{R}_{il}^{high} \\ \mathcal{G}_{il}^{high} \\ \mathcal{B}_{il}^{high} \end{bmatrix} ,$$

which is seen to be the desired generalized plane, hence only solutions fulfilling (5) are possible. It should be noted that *a)* since the mapping is linear it fits nicely into a least squares setting and *b)* that in such a least squares optimization the original 16th pixel does not get 'special treatment'. Thus this new parameterizations nicely solves the problem of spectral consistency.

IV. A More Sophisticated Image Prior

As mentioned above, using the quadratic error function, ρ_2 in the smoothness constraint, gives a rather blurry image. The reason being that it smooths across significant edges in the image. Here we address this problem by using non-linear diffusion or smoothing. That is, if there is a large support for an edge in the data smoothing should not be carried out across it. This has been shown to be equivalent to using robust error functions, ρ_{rob} instead of the 2-norm, ρ_2, c.f. [3]. Examples of such robust error functions include the truncated quadratic and the Huber norm.

These robust error functions are often implemented as weighted versions of the least squares norm, i.e.

$$\sum_{ij} \sum_{k \in \mathcal{N}_{ij}} \rho_2(\epsilon_{ijk}) \cdot w_{ijk} \tag{10}$$

where the weights are often updated in an iterative manner, giving rise to the iterative reweighted least squares (IRLS) type algorithms. An intuitive version of this reweighting is: that the algorithm should detect the significant lines in the image — over which the weights should be lowered — but that these significant lines are large gradients in the solution to the problem. A solution which in turn depends on the weights. *In this particular case reweighting is not required.* This is because we already have a high resolution version of the solution sought, namely the pan-chromatic image. So a main contribution of this paper is; proposing to get the weights, w_{ijk}, by analyzing gradients in the pan-chromatic image, as opposed to the solution $(\mathbf{R}^{high}, \mathbf{G}^{high}, \mathbf{B}^{high})$.

In this ongoing work, we have determined the weights as follows:

1) Apply the Canny edge-detector to the pan-chromatic image. [4]
2) Set the weights to $w_{ijk} = 0$ if there is an edge or else to $w_{ijk} = 1$.

This is a simple scheme, but serves well as a proof of concept, as seen in Section VI. The results, however, show that sometimes there is a need for more smooth transitions e.g. in that some less distinct edges are missed. This indicates that the weights should vary continuously between 0 and 1. Thus improvements on the weighting scheme are called for.

V. Proposed Algorithm

The proposed algorithm is thus as follows:

1) Get initial estimates of $(\mathbf{R}^{high}, \mathbf{G}^{high}, \mathbf{B}^{high})$ by use of (4), i.e. no smoothing applied.
2) Derive values for $(\mathcal{R}^{high}, \mathcal{G}^{high}, \mathcal{B}^{high})$ from $(\mathbf{R}^{high}, \mathbf{G}^{high}, \mathbf{B}^{high})$. Note that the values obtained from (4) are spectrally consistent, so this calculation is straightforward.
3) Derive the weights w_{ijk}, as described in Section IV.
4) Find the solution to the modified error function or

---

[4]The image is shifted half a pixel to fit with the borders of the pan-chromatic pixels.

objective function, now over $(\mathcal{R}^{high}, \mathcal{G}^{high}, \mathcal{B}^{high})$,

$$\min_{\mathcal{R}^{high}, \mathcal{G}^{high}, \mathcal{B}^{high}} \sum_{ij} \left(\mathcal{D}_{ij} + \gamma \sum_{k \in \mathcal{N}_{ij}} \rho_2(\epsilon_{ijk}) \cdot w_{ijk} \right) \tag{11}$$

5) Calculate the desired solution $(\mathbf{R}^{high}, \mathbf{G}^{high}, \mathbf{B}^{high})$ from $(\mathcal{R}^{high}, \mathcal{G}^{high}, \mathcal{B}^{high})$.

An additional advantage of the proposed approach is that (11) is a least squares function in $(\mathcal{R}^{high}, \mathcal{G}^{high}, \mathcal{B}^{high})$. Hence (11) is convex, implying that there are no local minima. The number of parameters of the problem, however, makes using the normal equations prohibitive. As part of our ongoing work, we are looking at how to exploit the structure of (11) to construct an efficient optimization algorithm.

VI. EXPERIMENTAL RESULTS

To validate the proposed approach we applied it to an image pair of Reykjavik, Iceland, acquired by the IKONOS earth imaging satellite on the 9th of August 2001. Two illustrative patches are depicted in Figure 1 and Figure 2. Here it is seen that the results are significantly improved, and that the final result clearly support the proposed algorithm. Although as stated above, there is room for improvement in the weighting scheme.

Fig. 1. A sample of the images: **top left:** Pan-chromatic image. **top right:** Extracted lines. **bottom left:** Initial estimates via (4) **bottom right:** Fused image via (11). The colors in the bottom row correspond to the red green and blue channel respectively.

VII. DISCUSSION AND FUTURE WORK

Here improvements to our satellite image fusion framework is presented, a framework purely based in the sensor physics and on prior assumptions on the fused image. The improvements ensure spectral consistency via reparameterizing, and better image priors by allowing breaks in the image smoothing where there is support for it in the image data.

As mentioned this is still ongoing work, and at present the main focus is on setting the weights better. The
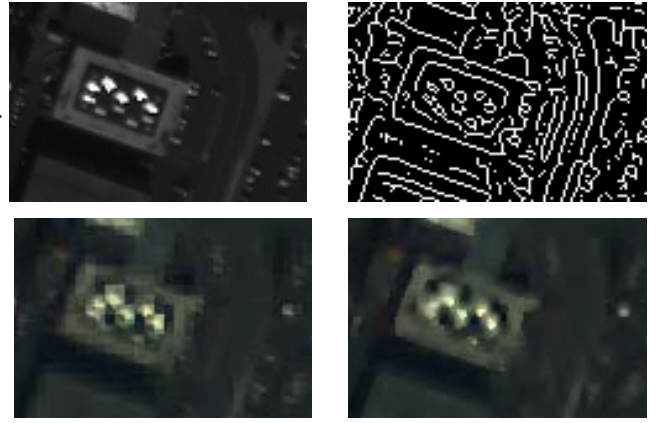

Fig. 2. Another sample of the images: **top left:** Pan-chromatic image. **top right:** Extracted lines. **bottom left:** Initial estimates via (4) **bottom right:** Fused image via (11). The colors in the bottom row correspond to the red green and blue channel respectively.

Canny edge detector and the binary weights do a — in our opinion — impressive job, but with room for improvements. It seems unnatural that the weights should not be continuous between 0 and 1 depending on the 'edginess' of the pan-chromatic image. The question is thus what measure of 'edginess' should be used?

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. *Geoscience and Remote Sensing, IEEE Transactions on*, 40(10):2300–2312, 2002.

[2] M. Beaulieu, S. Foucher, and L. Gagnon. Multi-spectral image resolution refinement using stationary wavelet transform. *IGARSS '03.*, 6:4032–4034, 2003.

[3] M.J. Black and A. Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *International Journal of Computer Vision*, 19(1):57–91, 1996.

[4] C. K. Munechika, J. S. Warnick, C. Salvaggio, and J. R. Schott. Resolution enhancement of multispectral image data to improve classification accuracy. *Photogrammetric Engineering and Remote Sensing*, 59(1):67–72, 1993.

[5] X. Otazu, M. Gonzalez-Audicana, O. Fors, and J. Nunez. Introduction of sensor spectral response into image fusion methods. application to wavelet-based methods. *Geoscience and Remote Sensing, IEEE Transactions on*, 43(10):2376–2385, 2005.

[6] C. Pohl and J.L. Van Genderen. Multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing*, 19(5):823–854, 1998.

[7] T. Ranchin and L. Wald. Fusion of high spatial and spectral resolution images: The arsis concept and its implementation. *Photogrammetric Engineering and Remote Sensing*, 66(1):49–61, 2000.

[8] A. Vesteinsson, H. Aanæs, J.R. Sveinsson, and J.A. Benediktsson. Spectral consistent satellite image fusion: using a high resolution panchromatic and low resolution multi-spectral images. *IGARSS '05.*, 4:2834–2837, 2005.