The curvelet transform for fusion of very-high resolution multispectral and panchromatic images

L. Alparone
Department of Electronics & Telecommunications, University of Florence, Florence, Italy

S. Baronti
Institute of Applied Physics “Nello Carrara” of the National Research Council IFAC–CNR, Florence, Italy

A. Garzelli & F. Nencini
Department of Information Engineering, University of Siena, Siena, Italy

Keywords: curvelet transform, image fusion, pan-sharpening, ridgelet transform

ABSTRACT: This paper presents a novel image fusion method, suitable for pan-sharpening of multispectral (MS) bands, based on multi-resolution analysis (MRA). The low-resolution MS bands are sharpened by injecting high-pass directional details extracted from the high-resolution panchromatic (Pan) image by means of the curvelet transform, which is a non-separable MRA, whose basis function are directional edges with progressively increasing resolution. The advantage with respect to conventional separable MRA, either decimated or not, is twofold: directional detail coefficients matching image edges may be preliminarily soft-thresholded to achieve denoising better than in the separable wavelet domain; modeling of the relationships between high-resolution detail coefficients of MS bands and of the Pan image is more fitting, being carried out in a directional wavelet domain. Experiments carried out on a very-high resolution MS + Pan QuickBird image show that the proposed curvelet method quantitatively outperforms state-of-the-art image fusion methods, in terms of geometric, radiometric, and spectral fidelity.

1 INTRODUCTION

Remote sensing image fusion aims at integrating the information conveyed by data acquired with different spatial and spectral resolutions, for purposes of photoanalysis, features extraction, modeling, and classification. A notable application is merging multi-spectral (MS) and panchromatic (Pan) images collected from space. Image fusion techniques may take advantage of the complementary spatial/spectral resolution characteristics for producing spatially enhanced MS observations.

Injection in the MS images of spatial details extracted from the high-resolution Pan image has been found to preserve the spectral characteristics of the original MS images. Multiresolution analysis, based on wavelets, filter banks, and Laplacian pyramids, has been recognised as a very efficient tools to implement fusion of images of different resolutions (G. Piella 2003). Data fusion methods, based on injecting high frequency components taken from the Pan image into resampled versions of the MS data, have demonstrated superior performances (Núñez et al. 1999; Aiazzi et al. 2002; Ranchin et al. 2003), especially concerning preservation of spectral information.

Redundant multi-resolution structures, like the generalised Laplacian pyramid (GLP) (Aiazzi et al. 2002), the undecimated discrete wavelet transform (UDWT) (Li et al. 2002) and the “à trous” wavelet transform (ATWT) (Núñez et al. 1999) have been found to be suitable for image fusion. Such decompositions are not critically subsampled; thus, inaccuracies in the fused images like ringing effects and canvas-like patterns originated by aliasing, are avoided (Gonzáles Audícana et al. 2004).

In very recent years, non-separable multi-resolution analysis has been introduced for image processing. In particular, the curvelet transform, proposed for image denoising by Starck et al. 2002; Candes et al. 2000; Donoho & Kodell 1994; Donoho 1995; 1992; Starck et al. 1998; 1998; 2001; 2000; 2002; 2000). The curvelet transform is a non-separable multi-resolution analysis tool that provides a directional multi-scale decomposition of images.
Data fusion based on multi-resolution analysis, however, requires the definition of a model establishing how the missing high-pass information to be injected into the resampled MS bands is extracted from the Pan image. This model is referred to in the literature (Garzelli & Nencini 2005; Ranchin et al. 2003) as inter band structure model (IBSM) and deals with the radiometric transformation (gain and offset) of spatial structures (edges and textures) when passing from Pan to MS images. The model is usually space-varying: it is calculated at a coarser resolution and inferred to the finer resolution. However, in order to increase it specificity, it would be desirable that such a model were calculated in a different domain, in which the linear structures that are injected were represented by few sparse coefficients.

In this work, we propose an image fusion method for Pan-sharpening of very-high resolution multi-spectral images, which operates in the non-separable transformed domain of curvelet transform coefficients. The algorithm is defined for either QuickBird or Ikonos-2 imagery, having scale ratio between Pan and MS equal to 4, but may be straightforwardly extended to other scale ratios. A performance comparison, carried out among advanced methods in the literature, on spatially degraded QuickBird data, whose reference originals are available, highlights the benefits of the proposed method for Pan-sharpening of satellite remote sensing imagery.

2 IMAGE FUSION BASED ON THE CURVELET TRANSFORM

In this section, firstly a multi-resolution structure suitable for image fusion is described. Its main feature is sensitiveness to directional edges and capability of representing object contours with few sparse nonzero coefficients. Secondly, we present the flowchart of a scheme suitable for Pan-sharpening of MS images, based on the curvelet transform and discuss the development of an IBSM in the curvelet domain.

2.1 Ridgelet transform

The first step is finding a transformation capable of representing straight edges with different slopes and orientations. The solution is the ridgelet transform (Do & Vetterli 2003), which may be regarded as the 1D wavelet transform of the Radon transform. However, an inconvenience with the ridgelet transform is that it is not capable of representing curve lines. To overcome this drawback, the input image is partitioned into square blocks and the ridgelet transform is applied to each block. Assuming a piecewise linear model for the contour, each block will contain straight edges only, that may be analyzed by the ridgelet transform. Figure 1 outlines the block ridgelet transform and highlights that the discrete Radon transform is obtained with a recto-polar resampling of the 2D-FFT of the image block.

2.2 Curvelet transform

The main benefit of curvelets is its capability of representing a curve as a set of superimposed functions of various lengths and widths. The curvelet transform, unlike wavelet transform, is a multi-scale transform, but, unlike wavelets, contains directional elements. Curvelets are based on multi-scale ridgelets with a bandpass filtering to separate image into disjoint scales. The side length of the localizing windows is doubled at every other dyadic sub-band. In practical applications, the discrete curvelet transform consists of applying the block (discrete) ridgelet transform in Fig. 1 to the detail frames of the ATWT. The algorithm (Starck et al. 2002) is outlined in the following:

- Apply the ATWT algorithm with $J$ scales. This transform decomposes an image $f$ in its coarse version, $c_J$, and in the details $\{d_j\}_{j=1,...,J}$ at scale $2^{-j}$.
\[ f(m, n) = c_J(m, n) + \sum_{j=1}^{J} d_j(m, n) \]  \hspace{1cm} (1)

- Select the minimum dimension of window, \( Q_{\text{min}} \), to apply to the finest scale \( d_1 \);
- For a fixed scale \( j \), make a partition of \( d_j \) in disjoint blocks having size

\[
Q_j = \begin{cases} 
2^j Q_{\text{min}} & \text{if } j \text{ is even} \\
2^{j-1} Q_{\text{min}} & \text{if } j \text{ is odd}
\end{cases}
\]  \hspace{1cm} (2)

- Apply the ridgelet transform to each block.

According to the original definition of curvelet transform, the base-band approximation \( c_J(m, n) \) is not further analysed by the block ridgelet transform shown in Fig. 1. However, this has been done in the present work, in order to derive the injection model in the directional domain.

2.3 Curvelet-based MS + Pan image fusion scheme

Figure 2 shows the flowchart of a CT-based scheme suitable for fusion of MS and Pan data, whose scale ratio is 4. Given one Pan image having finer resolution, and the \( L \) spectral bands of an MS image, having resolution coarser by 4, the goal is to obtain a set of \( L \) bands each having same spatial resolution as Pan. The enhancement of each band to yield the spatial resolution of Pan is synthesised from the levels \( s_1 \) and \( s_2 \) of the CT of the Pan image. The MS bands are preliminarily interpolated by \( p \) to match the scale of the Pan image. They constitute the low-pass component to which details extracted from Pan by means of CT are added. The two sets of CT coefficients, one for each layer, calculated from Pan are soft-thresholded to reduce the noise, weighted by the IBSM, and used to synthesize, by means of the inverse ridgelet transform, two maps of zero-mean spatial edges and textures that are added to the corresponding detail frames of the ATWT of the resampled MS bands. The Pan-sharpened MS image is obtained by ATWT synthesis, i.e. by summing approximations and enhanced detail frames of each band.
### 2.4 Injection model

The injection model is stated as an adaptive cross-gain, i.e. a ratio of local standard deviations. Such a gain weights CT coefficients, after they have been soft-thresholded to reduce the noise. In fact, the spatially uncorrelated noise is uniformly spread onto CT coefficients, unlike spatial details, that are concentrated in few sparse coefficients. In order to adaptively weight thresholded CT coefficients coming from high \( (s_1) \) and middle \( (s_2) \) resolution layers, the cross-gain must be calculated in a domain defined exactly on the same coordinates, that is, spatial block position and scale-angle pair within the block. Therefore, the approximations \( c_2 \) of MS bands and Pan are further analyzed with a block ridgelet transform, as shown in Fig. 1. As two block sizes exist for fine and coarse details, respectively, the ridgelet transform is applied to replicas of \( c_2 \) from MS and Pan, with different block sizes. Thus, there is a 1:1 correspondence between cross-gain and curvelet coefficient from Pan on both resolution layers \( (s_1) \) and \( (s_2) \).

### 3 EXPERIMENTAL RESULTS

#### 3.1 Quality assessment of fusion products

Quality assessment of Pan-sharpened MS images is a hard task (Chavez Jr et al. 1991; Wald et al. 1997; Piella & Heijmans 2003). Even when spatially degraded MS images are processed for Pan sharpening, and therefore reference MS images are available for comparisons, assessment of quality, or more exactly of fidelity to reference, usually requires computation of a number of different score indexes. Examples are correlation coefficient (CC) between each band of the fused and reference MS images, bias in the mean, root mean square error (RMSE), and spectral angle mapper (SAM), which measures the spectral distortion introduced by the fusion process.

Wald et al. (1997) proposed an error index that offers a global picture of the quality of a fused product. This error is called ERGAS, after its name in French, which means relative dimensionless global error in synthesis, and is given by:

\[
ERGAS = 100 \frac{1}{T} \sqrt{\frac{1}{K} \sum_{k=1}^{K} \left( \frac{\text{RMSE}(k)}{\mu(k)} \right)^2}
\]  

#### Figure 2. Flowchart of curvelet-based fusion of MS and Pan data with 1:4 scale ratio.
The curvelet transform for fusion of very-high resolution multi-spectral

where $h/l$ is the ratio between pixel sizes of Pan and MS, e.g. 1/4 for IKONOS-2 and QuickBird data, $\mu(k)$ is the mean (average) of the $k$th band, and $K$ is the number of bands. The score index (3) is a notable effort to encapsulate several measurements in a unique number. However, it does not consider CCs and especially fails in measuring spectral distortion, in addition to radiometric distortion.

An image quality index, namely Q4, for MS images having four spectral bands was recently proposed by the authors (Alparone et al. 2004) to be applied to Pan-sharpened MS images in order to assess the performance of different image fusion algorithms.

3.2 Performance comparison of image fusion algorithms

The proposed curvelet-based fusion procedure has been assessed on very high-resolution image data collected on June 23 2002 at 10:25:59 GMT+2 by the QuickBird spaceborne MS scanner on the urban and suburban areas of Pavia, Italy. The four MS bands span the visible and near infrared (NIR) wavelengths and are spectrally disjointed. The panchromatic band embraces the whole interval 450 ÷ 900 nm. All data are resampled to uniform ground resolutions of 2.8 m and 0.7 m GSD for MS and Pan. The full scale of all the bands is 2047 (11 bits) and is reached in the NIR wave-lengths. To obtain quantitative distortion measures, the Pan and MS bands were low-pass filtered and decimated by 4, to yield 2.8 m P and 11.2 m MS, respectively, and used to synthesize the four spectral bands back at 2.8 m. Thus, the true $512 \times 512$ MS data at 2.8 m are available for objective distortion measurements.

A performance comparison was carried out among the novel method and a number of state-of-the-art image fusion methods, which are listed in the following.

- Multiresolution IHS (Núñez et al. 1999) with additive model (AWL), based on ATWT.
- ATWT method with spectral distortion minimizing (SDM) model (Garzelli & Nencini 2005).
- ATWT method with Ranchin-Wald-Mangolini (RWM) model (Ranchin et al. 2003).
- ATWT method with context-based decision (CBD) model (Aiazzi et al. 2002).
- Gram-Schmidt spectral sharpening method (Laben & Brower 2000), implemented in ENVI®.
- High-Pass Filtering (HPF) technique (Chavez Jr et al. 1991), with $5 \times 5$ box filter for 1:4 ratio.

Also the case in which the MS data are simply resampled (through the 23-taps filter) and no addition of details is made, is presented to discuss the behaviors of the different fusion methods.

Table 1. Average cumulative quality/distortion indexes between 2.8 m MS spectral vectors and those obtained from fusion of 11.2 m MS with 2.8 m Pan. EXP denotes plain resampling.

<table>
<thead>
<tr>
<th></th>
<th>EXP</th>
<th>CT</th>
<th>AWL</th>
<th>SDM</th>
<th>RWM</th>
<th>CBD</th>
<th>GS</th>
<th>HPF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q4</td>
<td>0.750</td>
<td>0.888</td>
<td>0.827</td>
<td>0.864</td>
<td>0.865</td>
<td>0.881</td>
<td>0.874</td>
<td>0.857</td>
</tr>
<tr>
<td>SAM (deg.)</td>
<td>2.14°</td>
<td>1.76°</td>
<td>2.59°</td>
<td>2.14°</td>
<td>2.07°</td>
<td>1.85°</td>
<td>1.82°</td>
<td>2.33°</td>
</tr>
<tr>
<td>ERGAS</td>
<td>1.760</td>
<td>1.281</td>
<td>1.611</td>
<td>1.676</td>
<td>1.694</td>
<td>1.430</td>
<td>1.388</td>
<td>1.904</td>
</tr>
</tbody>
</table>

The parameters measuring the global distortion of pixel vectors, either radiometric (ERGAS) or spectral (SAM), and both radiometric and spectral (Q4) will give a comprehensive measure of quality, also matched by visual analysis. CT attains global scores better than those of the other methods, followed by GS, which is unknown in the literature, being patented (Laben & Brower 2000). Not surprisingly the SAM attained by CBD and RWM is lower than that of SDM (identical to that of resampled MS data), thanks to the unmixing capabilities of the former ones compared to the latter. The ranking of methods confirms that HPF is spectrally better than AWL(lower SAM), but radiometrically poorer (higher ERGAS). The novel index Q4 (Alparone et al. 2004) trades off both types of distortion and yields a unique quality index, according to which AWL is better than HPF.

Eventually, CT is compared with GS and with AWL on 0.7 m fusion products, as it happens in practice. Figure 3 displays the resampled 2.8 m MS bands and the enhanced bands at 0.7 m. A
visual inspection highlights that all the spectral signatures of the original MS data are carefully incorporated in the sharpened bands. AWL is geometrically rich and detailed, but unlikely over-enhanced. CT and GS products are visually superior and substantially similar to each other. Given the extremely different origin of the two algorithms (GS exploits a kind of component substitution, CT a non-separable multi-resolution analysis) and the fact that the GS algorithm is likely to have been strongly optimised for commercial exploitation, we are lead to draw the preliminary conclusion that the novel CT fusion method is promising for Pan-sharpening of MS images.

4 CONCLUSIONS

An original image fusion method based on the non-separable multi-resolution analysis provided by the curvelet transform and suitable for Pan-sharpening of MS bands, has been described and
assessed. The low-resolution MS bands are resampled to the fine scale of the panchromatic (Pan) image and sharpened by injecting high-pass directional details extracted from the high-resolution panchromatic (Pan) image. Thanks to the curvelet basis function, which are directional edges with progressively increasing resolution, a superior spatial enhancement is provided. Curvelet coefficients matching image edges are preliminarily soft-thresholded to achieve denoising better than in the separable wavelet domain (Starck et al. 2002). This issue is important, e.g. for fusion of either MS and synthetic aperture radar (SAR) data (Alparone et al. 2004b), or of Pan and hyperspectral data (Eismann & Hardie 2005). The second advantage of performing fusion in the curvelet domain is that modeling of the relationships between high-resolution detail coefficients of MS bands and of the Pan image is more fitting, being carried out in a directional multiresolution domain. Experiments carried out on a veryhigh resolution MS + Pan QuickBird image have demonstrated that the proposed curvelet method quantitatively outperforms state-of-the-art image fusion methods, in terms of geometric, radiometric and spectral fidelity.

REFERENCES


