Effects of Aliasing and Mis-Registration on Pan-Sharpening Methods Based on Either Component Substitution or Multi-Resolution Analysis

Andrea GARZELLI\textsuperscript{a1}, Bruno ALIAZZI\textsuperscript{b}, Luciano ALPARONE\textsuperscript{c}, Stefano BARONTI\textsuperscript{b}, and Massimo SELVA\textsuperscript{b}

\textsuperscript{a} DII, University of Siena, Italy
\textsuperscript{b} IFAC-CNR, Florence, Italy
\textsuperscript{c} DET, University of Florence, Italy

\textbf{Abstract.} In this paper, the characteristics of multi-spectral (MS) and panchromatic (P) image fusion, or pan-sharpening, methods, will be investigated. Depending on the way spatial details are extracted from P, such methods can be broadly labelled into two main classes, roughly corresponding to component substitution (CS), also known as projection substitution, and methods based on multi-resolution analysis (MRA), i.e. on digital filtering. Experimental results carried out on QuickBird and Ikonos data sets evidence that CS-based fusion is far less sensitive than MRA-based fusion to: a) registration errors, i.e. spatial misalignments between MS and P images, possibly originated by cartographic projection and resampling of individual data sets; b) aliasing occurring in MS bands and stemming from a modulation transfer function (MTF) of each MS channels that is excessively broad relative to the spatial sampling interval. Simulated misalignments, carried out at full scale by means of a suitable quality evaluation protocol, have evidenced the quality-shift trade-off of the two classes: MRA methods yield a slightly superior quality in the absence of misalignments, but are more penalized, whenever shifts between MS and P are present, than CS methods, which produce a slightly lower quality in the ideal case, but are intrinsically more shift tolerant.

\textbf{Keywords:} Aliasing, misregistration, multispectral imagery, pansharpening.

\textbf{Introduction}

Pan-sharpening is a branch of data fusion, more specifically of image fusion, that is receiving an ever increasing attention from the remote sensing community. New-generation space-borne imaging instruments operating in a variety of ground scales and spectral bands provide huge volumes of data having complementary spatial and spectral resolutions. Constraints on the signal to noise ratio (SNR) impose that the spatial resolution must be lower, if the desired spectral resolution is higher. Conversely, the highest spatial resolution is obtained whenever no spectral diversity is required. The trade-off of spectral and spatial resolution makes it desirable to perform a spatial enhancement of the lower resolution multi-spectral (MS) data or, equivalently, to increase the spectral resolution of the data-set having a higher ground resolution, but a lower spectral resolution; as a limit case, constituted by a unique panchromatic image (P) bearing no spectral diversity.

An extensive number of methods have been proposed in the literature over the last two decades. Most of them follow a general protocol, that can be summarised in the following two key points:

- extract high-resolution geometrical information of the scene, not present in the MS image, from the P image;

\textsuperscript{1} Corresponding author: Department of Information Engineering, University of Siena, Via Roma, 56 – 5311 Siena, Italy; Email: andrea.garzelli@unisi.it
incorporate such spatial details into the low-resolution MS bands, interpolated to the spatial scale of the P image, by properly modelling the relationships between the MS bands and the P image.

According to the most recent literature [1], pan-sharpening methods can be labelled into three categories.

The first category consists of methods based on multi-resolution analysis (MRA). They employ digital spatial filters to extract the high spatial frequency information from the P image. Such an information, constituted by high-pass spatial frequency component at each pixel, is possibly weighted by means of a suitable model [2] and added to the up-scaled MS images. Wavelets [3,4] and Laplacian pyramids [5] are MRAs most widely used in fusion methods.

The second category consists of methods based on component substitution (CS), also known as projection substitution. These methods do not involve any spatial filtering process. They make use of a spectral transformation to obtain a new projection of the MS and P images in which fusion occurs as substitution of one component with the P image. The inverse transformation produces MS images at the desired high resolution. Examples of spectral transformations are generalised Intensity-Hue-Saturation [6,7], Principal Component Analysis and Gram-Schmidt orthonormalisation procedure [8,9].

The third type of fusion methods makes use of both CS and MRA. The methods presented in [10-12] are examples of this type of hybrid fusion. In the case of [12], the spatial filtering of the intensity component and of the P image is designed and accomplished in the Fourier domain. Despite the hybrid nature of such methods, their behaviour is more similar to MRA methods than to CS methods. In many cases they are equivalent to an MRA fusion method with a specific injection model [13]. As an example, the extension of the additive wavelet luminance (AWL) method [10] to more than three bands constitutes the AWL proportional (AWLP) method [4], in which details extracted from P through MRA are injected proportionally to the original spectral vector, in order to preserve the spectral angle between the resampled original data and the fused data.

So far, the problem of how the various fusion methods behave in the presence of the two main types of impairments occurring in very high resolution (VHR) MS + P product, namely aliasing and mis-registration, has never been explicitly addressed. Users remark that CS methods are preferred because unavoidable registration inaccuracies of geocoded products and aliasing patterns, especially noticeable on sharp straight oblique contours, are mitigated after fusion with CS methods. Conversely, MRA methods are carefully avoided in these critical cases, because aliasing and/or mis-registration may lead to annoying visual artefacts. In this paper, the effects of aliasing and mis-registration on pan-sharpened imagery are investigated, for both CS- and MRA-based fusion methods.

1. A critical review of fusion methods

According to the most recent studies carried our by the authors, the majority of image fusion methods can be divided into two main classes, which have opposite behaviours towards aliasing and mis-registration. Such classes uniquely differ in the way the spatial details are extracted from the P image.

- Techniques that employ linear space-invariant digital filtering of the P image to extract the spatial details, i.e. the geometrical information, that will be added to the MS bands [3]; all methods employing MRA belong to this class.
- Techniques that yield the spatial details as pixel difference between the P image and a non-zero-mean component obtained from a spectral transformation of the MS bands, without any spatial filtering of the former. They are equivalent to substitution of such a component with
the P image followed by reverse transformation to produce the sharpened MS bands [9], same as for plain, i.e. not hybrid, CS-based methods.

Regardless of how spatial details have been obtained, their injection into the interpolated MS bands may be weighed by suitable gains, different for each band, possibly space-varying, i.e. a different gain at each pixel. Algorithms based on context-adaptive, i.e. local, models generally perform better than models fitting each band globally [2]. A pixel-varying injection model is capable of defining fusion algorithms based on modulation [13], e.g. Brovey for the class of methods in Figure 1b and SFIM [14] for the methods outlined in Figure 1a.

The two classes of methods described above exhibit complementary spectral-spatial quality trade-off. Methods without spatial filtering provide fused images with high geometrical quality of spatial details, but with possible spectral impairments. Methods employing spatial filtering are spectrally accurate in general, but may be unsatisfactory in terms of spatial enhancement. However, the spectral combination of bands is optimized for spectral quality of pansharpened products and spatial filtering is optimized for spatial quality, the two categories of methods yield very similar results in terms of overall quality [2].

Figure 1 shows the flowcharts of the two approaches. In the former case, filtering is crucial: MTF filtering yields best results [15]. In the latter case, key point is the spectral transformation defined by the set weights \( \{w_k\} \) [9]. Detail-injection gains \( \{g_k\} \) are used in both cases. The MTF low-pass filter in Figure 1b is used only to calculate the spectral weights and not to directly produce the spatial details, like in Figure 1a.

![Figure 1. Flowchart of the two main pan-sharpening approaches: (a) Based on filtering the P image, or more generally on multi-resolution analysis; (b) based on a spectral combination of bands, without filtering the P image, or more generally on component/projection substitution.](image)

1.1. Component substitution

CS-based pan-sharpening is a typology of simple and fast techniques based on a spectral transformation of the original bands in a new vector space. Most widely used transformations are intensity-hue-saturation (IHS), principal components analysis (PCA) and Gram-Schmidt orthogonalization procedure [8,9]. IHS fusion technique, originally defined for three bands only, has been extended to an arbitrary number of spectral bands [7]. The rationale of CS fusion is that one of the transformed components (usually the first component or intensity, \( I \)) is substituted by the high-resolution panchromatic image, \( P \), before the inverse transformation is applied. To ensure global preservation of radiometry, \( P \) is histogram-matched to \( I \), in such a way that the histogram-matched sharpening \( P \), once degraded to the spatial resolution of \( I \), exhibits same global mean and variance as \( I \). However, since the histogram-matched \( P \) and \( I \) may not have the same local radiometry, spectral distortion, appearing as local colour changes in a composition of three bands at a time, may occur in pansharpened products. To mitigate local spectral distortion, \( I \) may be taken as a linear combination of the MS bands with weighting coef-
coefficients adjusted to the extents of overlap between the spectral response of each MS channel and that of the P. In principle, if the lowpass approximation of P, synthesized by combining the spectral channels, exactly matches the low-resolution version of P, spectral distortion does not occur [7,9].

It is noteworthy that the spectral transformation need not be explicitly calculated to perform fusion [9]. In principle, any nonsingular square matrix would define a possible transformation. However, fusion depends only on the first transformed component defined as a linear combination of the MS bands with weights \{w_k\}. Therefore, fusion methods labeled as projection substitution and relative spectral contribution [1] are pretty equivalent.

1.2. Multiresolution analysis

MRA-based techniques substantially split the spatial information of the MS bands and of the P image into a series of bandpass spatial frequency channels. The high frequency channels are inserted into the corresponding channels of the interpolated MS bands. The sharpened MS bands are synthesized from their new sets of spatial frequency channels. The “à trous” wavelet transform and the Laplacian pyramid are most widely used to perform the MRA [10,5]. In such cases, the zero-mean high-frequency spatial details are simply given as the difference between, and its lowpass filter version \(P_L\). Recent studies [15] have demonstrated that if the lowpass filter is designed in such a way that it matches the modulation transfer function (MTF) of the spectral channel in which details will be injected, the spatial enhancement provided by MRA techniques becomes comparable to that of CS techniques.

MRA-based techniques may be accommodated within the framework of ARSIS (Amélioration de la Résolution Spatiale par Injection de Structures) [3,16], originally employing the decimated discrete wavelet transform (DWT).

2. Experimental results

A set of experiments has been carried out on VHR data, QuickBird and IKONOS, to evidence the sensitiveness of fusion to aliasing and mis-registration. Two methods have been chosen as peculiar of the two classes described. The improved Gram-Schmidt spectral sharpening (GS+) [9] is a modified version of ENVI’s GS [8] with improved spectral quality, thanks to a least squares (LS) calculation of the spectral weights \(w_k\) defining the generalized intensity I. The second method is the generalized Laplacian pyramid (GLP) [5] with Gaussian-like reduction filter [15] and almost ideal 23-taps expansion filter. The injection model, \(g_k\), is global in both cases and equal to the covariance between interpolated \(k\)th MS band and either I or lowpass filtered P (\(P_L\)), divided by the variance of either I or \(P_L\) [2].

2.1. Aliasing

QuickBird data have been chosen for this set of experiments, since the amount of aliasing is related to the value of the bell-shaped MTF at Nyquist frequency, which is greater for QuickBird than for IKONOS. Figures 2a-d and 2e-h report two distinct experiments. In both cases the startup are aliasing-free P and aliased MS. Aliasing is due to insufficient sampling and is noticeable around sharp oblique contours of the scene as annoying jagged patterns. Such patterns appear in Figures 2b and f interpolated at the spatial scale of P. The outcomes of the two methods, GS+ and GLP are completely different: the former rejects aliasing almost totally; the latter leaves aliasing patterns almost unchanged. Spectral quality is always good, thanks to the good properties of all MRA methods in general and of GS+ in particular: no significant change in colour hues is perceivable in all fusion
products. Under the perspective of aliasing rejection, a CS method optimized for spectral quality [9] is undoubtedly preferable to an MRA-based method, yet optimized for spatial quality [15].

Table 1. Geometrical ($D_g$) and spectral (SAM) distortions of fusion based on filtering (GLP) and on plain projection (GS+). Test image IKONOS, 1024×1024. Interpolated MS and P exactly overlapped and misregistered (MIS) by 4 pels along $x$ and $y$.

<table>
<thead>
<tr>
<th></th>
<th>GLP</th>
<th>MIS GLP</th>
<th>GS+</th>
<th>MIS GS+</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_g$</td>
<td>0.0860</td>
<td>0.2869</td>
<td>0.0974</td>
<td>0.2017</td>
</tr>
<tr>
<td>SAM</td>
<td>0.4359</td>
<td>3.1315</td>
<td>0.4691</td>
<td>3.1562</td>
</tr>
</tbody>
</table>

Figure 2. Details of original and fused QuickBird data (256×256 at 0.7 m scale). (a),(e): panchromatic; (b),(f): true colour display of 2.8 m MS interpolated at 0.7 m; (c),(g): enhanced Gram-Schmidt; (d),(h): generalized Laplacian pyramid with global injection model.

2.2. Misregistration

An IKONOS dataset has been chosen for simulations, because of the low amount of aliasing, which could mask the effects of mis-registration. The MS and P data sets have been misaligned along both $x$ and $y$ by 4 pixels at P scale, i.e. one pixel at MS scale. The fusion methods compared are still GS+ and GLP. Two distinct experiments are shown in Figure 3. For each experiment, P, interpolated MS, GS+ fusion of overlapped MS and P, GS+ fusion of misregistered MS and P (still), GLP fusion of overlapped data and GLP fusion of misregistered data are displayed. The visual results are stunning. While GS+ and GLP behave quite similarly in absence of mis-registration, on misaligned data, the former produces an image with high resemblance to the ideal case, apart from colours, which are obviously shifted together with the MS original. The geometry of spatial details is preserved to a large extent notwithstanding the 4-pel shift. Conversely, GLP exhibits a better preservation of colours, but can not tolerate the 4-pel shift. In Figure 3j the circular target is split into its lowpass and edge components. In Figure 3l the geometry of the scene of roofs fades off.

In order to quantify the losses in quality, either spectral or spatial/geometric, of the two sample methods, GS+ and GLP, in the presence of mis-registration, separate measurements of spectral angle (SAM) and of spatial distortion, $D_g$, according to the QNR protocol [17] have been performed.
between fused and unfused MS data. The choice of distortion metrics stems from the need of carrying out measurements at full spatial scale. All distortions are calculated with the overlapped MS as reference, also when MS is shifted. Table 1 highlights that average SAM of the images obtained from misregistered data is simply achieved by adding the same constant offset to the corresponding distortions in the case of null shift. Conversely, GLP yields $D_S$ 12% lower than GS+ without misregistration; but when the data are shifted GS+ attains a distortion that is almost 30% lower than GLP. Such distortion values match the higher degree of shift-tolerance of GS+, visually remarked in Figure 3.

3. Conclusions

This work has pointed out the suitability and criticalness of pan-sharpening methods, by providing a general framework in which the almost totality of existing methods can be accommodated into two
classes, roughly matching the earlier subdivision in component substitution (or component projection) and methods based on multiresolution analysis (or digital filtering). The sole difference of methods from the two classes is the intrinsic modality of detail extraction from the panchromatic image. Whenever spatial details are extracted by highpass filtering \( P \), the spectral quality of fusion products is guaranteed even in the case of MS and P datasets acquired at different times. However, the spatial quality is extremely sensitive to spatial misalignments between MS and P, which lead to a progressive fading of the geometry of the scene, as well as to aliasing occurred in the digitization of the optical MS bands, which leads to the introduction of annoying jagged patterns in the pansharpened image. Conversely, the class of methods not employing digital spatial filters on the P image is extremely robust to mis-registration between interpolated MS and P and especially to aliasing in the MS image. Under general and likely assumptions, impairments due to either aliasing or mis-registration (also both together) are very little noticeable in fusion products, especially in terms of spatial quality.

References


