Data fusion: taking into account the modulation transfer function in ARSIS-based pansharpening methods

Pierre Massip, Philippe Blanc, Lucien Wald

Massip P., Blanc Ph., Wald L., 2010. Data fusion: taking into account the modulation transfer function in ARSIS-based pansharpening methods. In Proceedings of the 29th Symposium of the European Association of Remote Sensing Laboratories, Chania, Greece, 15-18 June 2009. I. Manakos and C. Kalaitzidis (Eds.). Published by IOS Press, 424-431, doi:10.3233/978-1-60750-494-8-320

Data fusion: taking into account the modulation transfer function in ARSIS-based pansharpening methods

Pierre MASSIP [a], Philippe BLANC [a] and Lucien WALD[a,1]
[a] *Center of Energy and Processes, Mines ParisTech /Armines, 06904 Sophia Antipolis, France*

Abstract. Multispectral images provided by satellite have a poor spatial resolution while panchromatic images (PAN) exhibit a spatial resolution two or four times better. Data fusion is a mean to synthesize MS images at higher spatial resolution than original by exploiting the high spatial resolution of the PAN. This process is often called pan-sharpening. The synthesized multispectral images should be as close as possible to those that would have been acquired by the corresponding sensors if they had this high resolution. The methods based on the concept "Amélioration de la Résolution Spatiale par Injection de Structures" (ARSIS) concept are able to deliver synthesized images with good spectral quality but whose geometrical quality can still be enhanced. We propose to consider the characteristics of the sensor to improve the geometrical quality. We take explicitly into account the modulation transfer function (MTF) of the sensor in the fusion process. Though this study is limited in methods and data, we observe a better restitution of the geometry and an improvement in the majority of quality indices classically used in pan-sharpening. The communication also presents a means to assess the respect of the synthesis property from a MTF point of view.

Keywords. Image fusion, pan-sharpening, multispectral, panchromatic, PSF, MTF, ARSIS, remote sensing

Introduction

Image fusion is a vivid topic in remote sensing, among other domains. Fusion was defined by the EARSeL Working Group [1] as "a formal framework in which are expressed means and tools for the alliance of data originating from different sources. It aims at obtaining information of a greater quality, although the exact definition of 'greater quality' will depend on the application". Here we specifically focus on the synthesis of multispectral (MS) images at a higher spatial resolution than original by exploiting the high spatial resolution of another image.

Current satellites provide images of two kinds with opposite characteristics. Multispectral (MS) images have a good spectral quality but a poor spatial resolution whereas panchromatic (PAN) images have on the contrary a high spatial resolution but with a poorer spectral quality. These two types of images allow to well identify observed structures through different information by geometrical information on the

[1] Corresponding Author: L. Wald, Mines ParisTech/Armines, Sophia Antipolis, 06904, France.
E-mail: Lucien.wald@mines-paristech.fr

one hand and by the spectral signature on the other hand. The interest of combining both types of images has been demonstrated by many authors.

Data fusion is a means to synthesize MS images at higher spatial resolution than original by exploiting the high spatial resolution of the PAN. This process is often called pan-sharpening. Here, the aim of fusion is to perform a high-quality transformation of the MS content when increasing the spatial resolution from the original one to that of the PAN image. The problem may be seen as the inference of the information that is missing to the original MS images for the construction of the MS images synthesized at a better resolution [5-7].

[2-4] establish that the fused images should obey several properties, that were synthesized by the EARSeL Working Group [6-7]. The consistency property expresses that the fused images should offer a strong consistency with the original data set: if a fused product is downsampled to its original low resolution, the original MS image should be retrieved. The synthesis property deals with the quality of the high-quality transformation: any synthetic image should be as identical as possible to the image that the corresponding sensor would observe with the highest spatial resolution, if existent. This holds for each individual modality MS and the multi-modality set.

Recent published works tend to demonstrate that the methods based on the ARSIS concept better synthesize images with respect to the expected spectral properties [2, 7-12]. However, the need to refine the spatial/geometrical quality of synthesized image is often underlined [2] [7] [12-13]. Geometry can be defined as the spatial organization of the radiometry in the image [14]. Characteristics of the imaging instrument are important with respect to the restitution of the geometry of the landscape. Besides its spectral and spatial resolutions, an important characteristic of optical instruments is the impulse response which is named point spread function (PSF). The PSF has an associate value, the modulation transfer function (MTF). The MTF is the modulus of the Fourier transform of the PSF. According to [15], the MTF determines the upper limits to the image quality, *i.e.*, the image resolution or sharpness and describes the image quality in terms of contrast as a function of spatial frequency. The difference in sampling rates between MS and PAN images implies a difference in the MTF of these images. This difference must be taken into account when synthesizing MS images at the PAN spatial resolution. A few fusion methods already include this difference in MTF through the multiscale model (MSM) in the ARSIS concept [2] [13] [16-18].

Our manner to treat the MTF is different from these previous works. We take explicitly into account the MTF of the sensor in the fusion process for each modality, independently from the MSM.

1. Objective

We now express the synthesis property in terms of geometry: the restitution of the radiometric organization of the observed landscape by the synthesized image should be as identical as possible to the organization that the corresponding sensor would observe with the highest spatial resolution, if existent. The work presented here aims at including the MTF into an existing fusion method in order to demonstrate that taking into account the difference in MTF leads to a better respect of the synthesis property from the geometrical point of view without degrading the quality of the other aspects of the synthesized images.

As said before, we selected an existing method belonging to the ARSIS concept. This method is called ATWT-M3, where ATWT denotes the MSM and means “à trous” wavelet transform, and M3 is the inter-band structure model described by [5], [10]. This method has been shown to provide good results in most cases; it already includes the difference in MTF in its MSM. Therefore, it is considered as an efficient method against which we will compare our own results. Whereas the results will not be the same than those obtained here, our way of taking into account the difference in MTF may easily apply to other methods belonging to the ARSIS concept and likely less easier, to other methods.

By explicitly integrating the difference in MTF, the aim is to give a 'real' MTF to the synthesized images. In figure 1, are drawn the MTFs for the four MS channels of the Pleiades sensor, discussed later, and that for the PAN, as a function of the frequency f normalized with the sampling frequency of the PAN. In this case, the spatial resolution of the PAN image (70 cm) is 4 times better than the MS (280 cm). Note that in this figure, the MTF of the MS are zero for relative frequencies equal to, or greater than, 0.25 (=1/4). It means that one cannot distinguish details less than 280 cm in size in the MS images. On the contrary, the MTF of PAN is equal to 0.4. The difference in MTF between low and high spatial resolution images is evidenced in this figure: the MTF of MS has a sharper decrease with frequency than that of PAN. Of particular interest here is the relative Nyquist frequency for MS, equal to 0.125. It denotes the smallest frequency that can be detected in MS images, i.e., twice the pixel size.

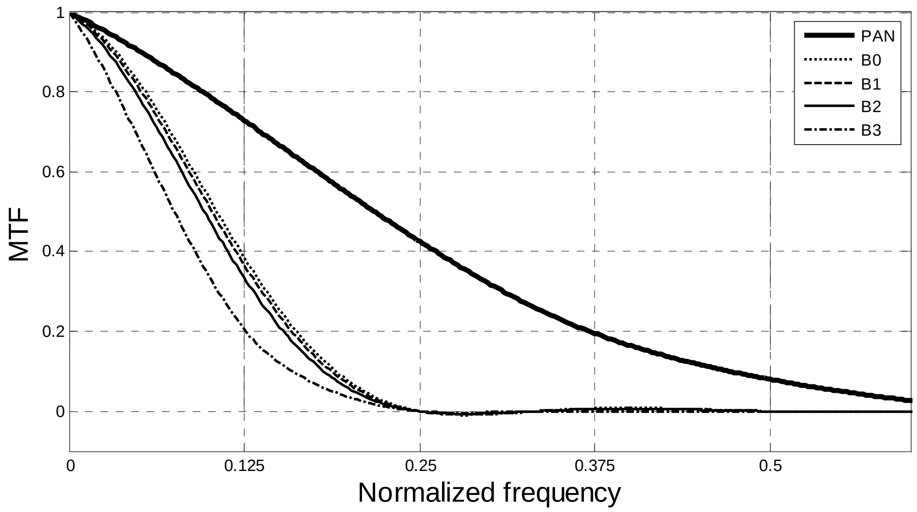


Figure 1. Modulation transfer function for MS and PAN images.

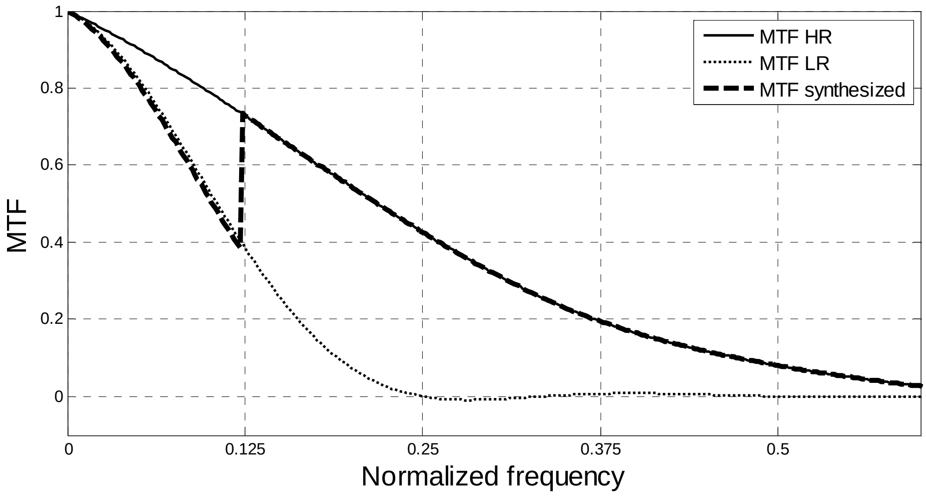


Figure 2. Schematic representation of the MTF of the synthesized MS images. LR and HR, respectively, mean low and high resolution.

When increasing the spatial resolution by injecting high-frequencies into the original MS images, the MTF of the synthesized image should look like that of PAN. However, without modification of the MTF during the fusion process, the resulting MTF exhibits a discontinuity for the relative frequency 0.125 as schematically drawn in Fig. 2. To attenuate the gap between the value of MS and PAN at this frequency, a solution consist in “raising” the MTF of the MS frequencies in the range [0, 0.125] so that it is close to that of a high resolution image. Doing so provides a MTF closer to a “real” MTF and similar contrast in the image for a same frequency without considering its origin: PAN or MS.

2. Data set

The data set used for this work is a set of images resulting from the simulation of the Pleiades mission of the Centre National d'Etudes Spatiales (CNES). These images are simulated from images at a higher resolution from the airborne sensor Pelican. CNES kindly provides us with MS images at both high (70 cm) and low (280 cm) spatial resolution and a PAN image at high resolution (70 cm). The tests have been done on an excerpt of the city of Toulouse in France.

These images result from simulation; it means that we have the so-called original MS images at 280 cm, the PAN image at 70 cm and the actual MS images at 70 cm which constitute the reference that should be attained. This is not the common situation in which the actual MS images at high resolution are unknown. However, it is a very convenient situation to assess the influence of a parameter such as the MTF because all parameters are known. During the creation of the simulated Pleiades images, the MTF is controlled by CNES for each image and resolution. Consequently, the MTF of MS, PAN and actual MS images are known for this study.

3. Modification of MTF for fusion in ARSIS based methods

As said before, we exploit an existing method: ATWT-M3, and we create a new one: ATWT-M3-FTMadapted by taking the MTF explicitly into account in the method. We transform the MTF before the decomposition of the MS image by the undecimated wavelet transform (“à trous” wavelet transform). The application of the undecimated transform implies a resampling of the original MS images to the high resolution. The modification of the MTF is done on this resampled MS image.

The transformation starts with a deconvolution of the original MS images at low resolution. The deconvolution kernel corresponds to the detector part of the MTF —the optical MTF is not taken into account.— Then, we perform the resampling of this image that has now only the optical MTF. The resampling is of spline type. Then we convolute the resampled images with the MTF of a high resolution detector.

After that, the method ATWT-M3 is applied which is based on undecimated wavelet transform [5]. For each modality, the high frequencies wavelet coefficients from the PAN image are transformed into MS high frequencies wavelet coefficients by applying an affine function. The parameters of this function are determined by a least-square fitting between the PAN and MS low-frequencies wavelet coefficients [5]. Finally, the inverse wavelet transform is applied to obtain the synthesized MS images at high resolution.

4. Results

The results of the two methods: ATWT-M3 and ATWT_M3_FTMadapted are compared to the reference image. As recommended by several authors [3-4] [6] [12], we perform statistical comparisons and a visual analysis of the possible discrepancies.

4.1. Statistical results

Quality budgets are obtained by the very same tool than that used for the "data fusion contest 2006" [12]. Table 1 exhibits the result of mono-modal indices: relative bias (biasRel), relative difference variance in percent (diffVarRel), relative standard deviation (σRel), correlation coefficient (cc), and correlation coefficient of high frequency (ccHF), as well multimodal indices: ERGAS and mean SAM.

The ideal values are presented for each index. This table shows that the bias for each method is very close to the ideal value; this is characteristic of the methods based on the ARSIS concept [5] [7] [12]. The bias is smaller for ATWT-M3 than for ATWT-M3-FTMadapted but in both cases, it remains very close to 0. For the other parameters, one notes a large increase in quality budget when adapting the MTF.

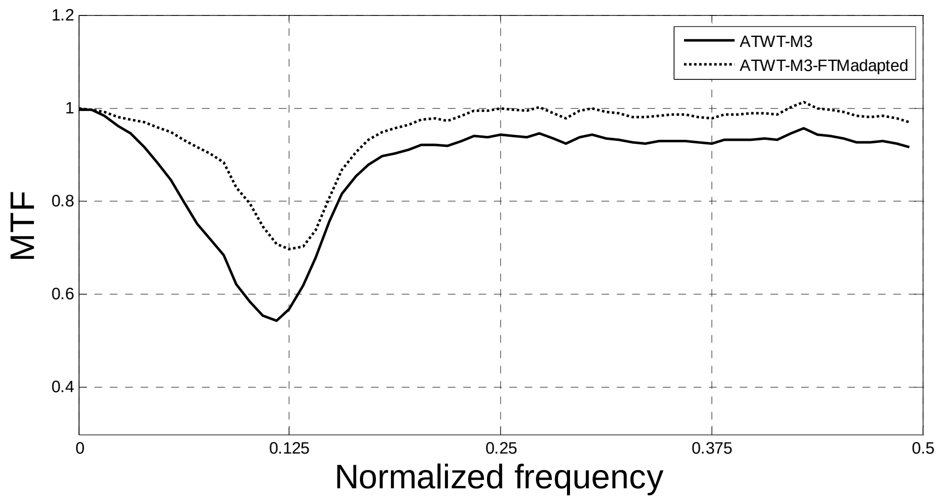
Table 1. Comparison of quality budget for ATWT-M3 and ATWT-M3-FTMadapted

modalities	Indexes	Ideal	ATWT-M3	ATWT-M3-FTMadapted	Delta in percent
bleu	biaisRel	0	-0,001	0,006	-670
	diffVarRel	0	14,4	3,3	77
	σRel	0	8,7	7,2	18
	Cc	1	0,9833	0,9875	0.4
	ccHF	1	0,9203	0,9224	0.2

green	biaisRel	0	-0,001	0,007	-1100
	diffVarRel	0	15,5	4,3	72
	σRel	0	8,6	6,3	26
	Cc	1	0,9875	0,9921	0.5
	ccHF	1	0,9605	0,9628	0.2
red	biaisRel	0	0,001	0,012	-1836
	diffVarRel	0	15,8	6,0	62
	σRel	0	9,3	6,0	35
	Cc	1	0,9905	0,9953	0.5
	ccHF	1	0,9690	0,9734	0.5
NIR	biaisRel	0	0,016	0,018	-11
	diffVarRel	0	34,2	26,7	22
	σRel	0	13,7	10,9	21
	Cc	1	0,9669	0,9793	1.3
	ccHF	1	0,8783	0,8886	1.2
multimodal indexes	ERGAS	0	2,57	1,96	24
	mean SAM	0	3,07	2,89	6

All these indices demonstrate improvements. Nevertheless, they do not inform on the aspect of the MTF of the synthesized images. Consequently, we perform a comparison of MTF between the reference and the synthesized images in order to evaluate if the modification of the MTF meets the expectations. This comparison is based on the cross correlation in the Fourier domain (of the PSD of the images) and is made relative to the MTF of the reference image:

$$mtf_{rel} = \frac{cross\ psd2d\,(xs_{ref}, xs^{*})}{psd\,(xs_{ref})} \tag{1}$$



(1)

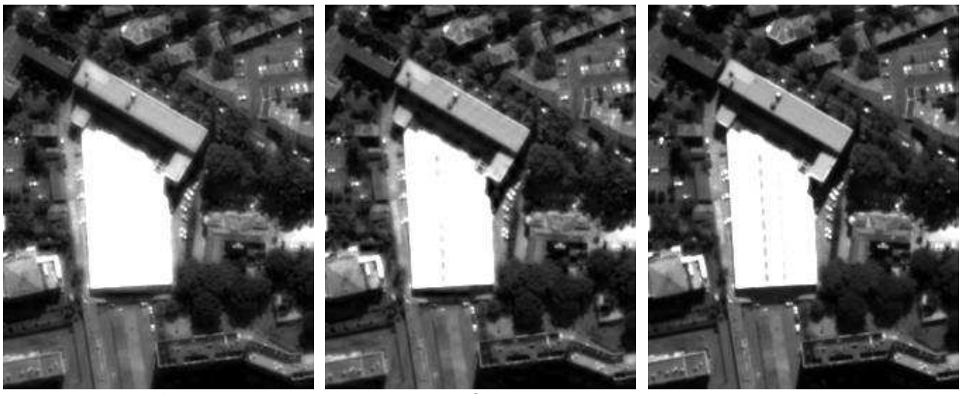
Figure 3. MTF of ATW-M3 and ATWT-3-FTM adapted relative to the MTF reference.

The MTF_{rel} denotes the behaviour of the MTF of the synthesized image compared to that of the reference image, as a function of the relative frequency (Fig. 3). The ideal result is a straight line for a value of one. Fig. 3 displays the results for band B2. corresponding to the wavelength of the red.

In this figure, the MTF_{rel} with adaptation of MTF (ATWT-M3-FTMadapted) is closer from the ideal value 1 than that of the method ATWT-M3, and this for all frequencies. Consequently, the improvement brought by the adaptation of the MTF is demonstrated in this case. One may note in this figure that the MTF_{rel} exhibits a trough between 0.1 and 0.15. It is believed that this is due to the fact that only the low frequency of the resampled image are injected in the synthesized image and also because information between the PAN and each MS band are not the same [10] which make that the PAN information is not fully exploited to create the high frequencies wavelet coefficients. This is under investigation and should be confirmed.

4.2. Visual analysis

Figure 4 exhibits three images: one without MTF adaptation (a), one with MTF adaptation (b) and the image of reference (c). In this excerpt we observe that the contour line of the white building is sharper in (b) than in (a) and that the image (b) is the closest to (c). This holds also for color composites and here again, we find an improvement when adapting the MTF.



a) b) c)

Figure 4. Excerpt of Pleiades images over Toulouse, France. Red modality synthesized with a) ATWT-M3, b) ATWT-M3-FTMadapted, c) reference image. Copyright CNES 2000.

5. Conclusion

We have shown how the difference in MTF between low and high resolution can be taken into account into a “classical” fusion method based on the ARSIS concept. In a given case, we have observed a better restitution of the geometry and an improvement in the majority of quality indices classically used in pan-sharpening. A new measure has been proposed to evaluate the respect of the synthesis property from a geometrical point of view.

Take into account the difference of MTF between MS and PAN images allow to have synthesized images of better quality. This work can be apply to others concept. It could be interesting to observe the difference of behavior of this technique depending on the concept it's applied on. The MTF relative that is given by our tool let us observe improvement and weakness of the MTF of the fuse image.

Though this study is limited in methods and data, the present results are encouraging and may constitute a new way to improve the restitution of geometrical features by already efficient fusion methods.

References

[1] L. Wald, Some terms of reference in data fusion. *IEEE Transactions on Geosciences and Remote Sensing*, 37 (1999), 1190-1193.

[2] T. Ranchin, B. Aiazzi, L. Alparone, S. Baronti, and L. Wald, Image fusion—The ARSIS concept and some successful implementation schemes. *ISPRS J. Photogramm. Remote Sens.*, vol. 58, no. 1/2, pp. 4–18, Jun. 2003.

[3] L. Wald, T. Ranchin, and M. Mangolini, Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogrammetric Engineering & Remote Sensing*, 63 (1997), 691-699.

[4] J. Li, Spatial quality evaluation of fusion of different resolution images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, 33 (2000), B2-2, 339–346.

[5] 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 (2000), 49-61.

[6] C. Thomas and L. Wald, Assessment of the quality of fused products. In Proc. 24th EARSeL Symp. New Strategies for Eur. Remote Sens., Dubrovnik, Croatia, Oluic, Ed. Rotterdam, The Netherlands: Millpress, May 25–27, 2004, pp. 317–325.

[7] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Transactions on Geosciences and Remote Sensing*, 46 (2008), 1301-1312.

[8] D. A. Yocky, Multiresolution wavelet decomposition image merger of Landsat Thematic Mapper and SPOT panchromatic data. *Photogramm. Eng. Remote Sens.*, 62 (1996), 1067–1074.

[9] Aiazzi B., L. Alparone, S. Baronti and A. Garzelli, Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Transactions on Geosciences and Remote Sensing*, 40 (2002), 2300-2312.

[10] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, A comparative analysis of image fusion methods. *IEEE Trans. Geosci. Remote Sens.*, 43 (2005), 81–84.

[11] P. Terretaz, Comparison of different methods to merge SPOT P and XS data : Evaluation in an urban area. In Proceedings, EARSeL Symposium 1997 " Future Trends in Remote Sensing," Lyngby, Denmark, P. Gudmansen Ed., A. A. Balkema Publ., 1998, 435-443.

[12] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data fusion contest. *IEEE Trans. Geosci. Remote Sens.*, 45 (2007), 3012–3021.

[13] C. Thomas, Fusion d'images de résolutions spatiales différentes (Fusion of images of different spatial resolutions). Ph.D Thesis, Ecole des Mines de Paris, Paris, France, 11 December 2006, 165 p. and annexes.

[14] G. Peyré, Géométrie multi-échelles pour les images et les textures. Report Ecole Polytechnique, Centre de Mathématiques Appliquées, Palaiseau, France, 2005.

[15] D. Grois, I. Shcherback, T. Danov, and O. Yadid-Pecht, Theoretical approach to CMOS APS PSF and MTF modeling – evaluation. Sensors Journal, IEEE 6 (2006), 118-124.

[16] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, M. Selva, A MTF-based spectral distortion minimizing model for pan-sharpening of very high resolution multispectral images of urban areas. In Proceedings 2nd GRSSDSPRS Joint Workshop on "Data Fusion and Remote Sensing over Urban Areas", 2003, 90-94.

[17] A. Garzelli, F. Nencini, Interband structure modeling for Pan-sharpening of very high-resolution multispectral images. *Information Fusion*, 6 (2005), 213 - 224

[18] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, M. Selva, MTF-tailored multiscale fusion of high-resolution MS and Pan imagery. *Photogrammetric Engineering & Remote Sensing*, 72 (2006), 591-596.