Comparison of Pansharpening Algorithms: Outcome of the 2006 GRS-S Data Fusion Contest

Luciano Alparone, Lucien Wald, Jocelyn Chanussot, Claire Thomas, Paolo Gamba, Lori-Man Bruce

HAL Id: hal-00177641

https://hal.archives-ouvertes.fr/hal-00177641

Submitted on 8 Oct 2007

Comparison of Pansharpening Algorithms: Outcome of the 2006 GRS-S Data-Fusion Contest

Luciano Alparone, Lucien Wald, Jocelyn Chanussot, *Senior Member, IEEE*, Claire Thomas,
Paolo Gamba, *Senior Member, IEEE*, and Lori Mann Bruce, *Senior Member, IEEE*

Abstract—In January 2006, the Data Fusion Committee of the IEEE Geoscience and Remote Sensing Society launched a public contest for pansharpening algorithms, which aimed to identify the ones that perform best. Seven research groups worldwide participated in the contest, testing eight algorithms following different philosophies [component substitution, multiresolution analysis (MRA), detail injection, etc.]. Several complete data sets from two different sensors, namely, QuickBird and simulated Pléiades, were delivered to all participants. The fusion results were collected and evaluated, both visually and objectively. Quantitative results of pansharpening were possible owing to the availability of reference originals obtained either by simulating the data collected from the satellite sensor by means of higher resolution data from an airborne platform, in the case of the Pléiades data, or by first degrading all the available data to a coarser resolution and saving the original as the reference, in the case of the QuickBird data. The evaluation results were presented during the special session on Data Fusion at the 2006 International Geoscience and Remote Sensing Symposium in Denver, and these are discussed in further detail in this paper. Two algorithms outperform all the others, the visual analysis being confirmed by the quantitative evaluation. These two methods share the same philosophy: they basically rely on MRA and employ adaptive models for the injection of high-pass details.

Index Terms—Image fusion, multispectral (MS) imagery, pansharpening, quality assessment, QuickBird (QB), simulated Pléiades data.

I. Introduction

DATA fusion is a formal framework for combining and utilizing data originating from different sources. It aims to obtain information of greater quality; the exact definition of "greater quality" will depend upon the application [1]. The application foreseen in this paper is the synthesis of multispectral (MS) images at a higher spatial resolution by exploiting an alternate high-resolution image acquired in another modality. These synthetic images should be similar to MS images that would have been observed with a sensor at higher resolution [2], [3]. When the higher resolution image is panchromatic (Pan), i.e., a single wide spectral-band image acquired across the visible and possibly near-infrared (NIR) wavelengths, this fusion process is usually called pansharpening of MS images.

Spaceborne sensors such as SPOT, Ikonos, or QuickBird (QB) provide images with different characteristics: on one hand, images with high spectral resolution but low spatial resolution and, on the other hand, images with low spectral resolution but high spatial resolution. Several works have demonstrated the usefulness of fused products offering high spectral and spatial resolutions at the same time in various environmental applications [4]–[11]. Pansharpened products are becoming very popular (for example, Google Earth), and data providers are offering higher and higher amounts of them at lower and lower costs.

A variety of pansharpening techniques take advantage of the complementary characteristics of spatial and spectral resolutions of the data [12]. Among them, component-substitution (CS) methods [13] are attractive because they are fast, easy to implement, and allow user's expectations to be fulfilled. When exactly three MS bands are concerned, the most widely used CS fusion method is based on the intensity–hue–saturation (IHS) transformation. The spectral bands are resampled and coregistered to the Pan image before the IHS transformation is applied. The smooth intensity component I is substituted with the high-resolution Pan and transformed back to the spectral domain via the inverse IHS transformation. This procedure is equivalent to inject, i.e., add, the difference between the sharp Pan and the smooth intensity I into the resampled MS bands [14]. Usually, the Pan image is histogram-matched, i.e., radiometrically transformed by a constant gain and bias in such a way that it similarly exhibits mean and variance as I, before substitution is carried out. However, since the histogram-matched Pan and the intensity component I do not generally have the same local mean, when the fusion result is displayed in color composition, large spectral distortion may be noticed as color changes. This effect occurs because the spectral response of the I channel, as synthesized by means of the MS bands, may be far different from that of Pan. Thus, not only spatial details but also slowly space-varying radiance offsets are locally injected. Generally speaking, if the spectral responses of the MS channels are not perfectly overlapped with the bandwidth of

Pan, as it happens with the most advanced very high resolution imagers, Ikonos and QuickBird, IHS-based methods may yield poor results in terms of spectral fidelity [15].

A viable solution is to define a generalized IHS (GIHS) transform by including the response of the NIR band into the intensity component [16]. The choice of the weights of the transformation can be related to the spectral responses of the Pan and MS bands as in [16] and [17], where prefixed weights are adopted. Thus, the spectral quality of fusion results may be notably improved. However, in order to definitely overcome the inconvenience of spectral distortion, methods based on injecting zero-mean high-pass spatial details extracted from the Pan image only have been extensively studied.

Multiresolution analysis (MRA) provides effective tools, like wavelets and Laplacian pyramids [2], [3], [18]–[20], to help carry out data-fusion tasks. However, in the case of high-pass detail injection, spatial distortions, typically, ringing or aliasing effects, originating shifts or blur of contours and textures, may occur in fusion products. Such impairments, which may be as much annoying as spectral distortions, are emphasized by misregistration between MS and Pan data, particularly if the MRA underlying detail injection is not shift-invariant [19], [21].

MRA-based fusion requires definition of a model establishing how the missing high-pass information is to be extracted from the Pan image and then injected into the MS bands. It may be accomplished either in the domain of approximations between each of the resampled MS bands and a low-pass version of the Pan image having the same spatial frequency content as the MS bands, or in the domain of medium frequency details, in both cases by measuring local matching [3], [22]. High-frequency details are not available for MS bands and must be inferred through the model, starting from those of Pan.

This paper is organized as follows. The motivations underlying the image fusion contest and the scope of its outcome are highlighted in Section II. Contest participants are presented in Section III, and their methods are briefly outlined. In Section IV, the data sets used throughout the context are described, together with the evaluation criteria that have been chosen. The steps of the evaluation procedure and some of the intermediate results are reported in Section V, together with the final ranking of the methods. Section VI contains a discussion that aims to explain the results based on the characteristics of the methods that performed best. Finally, conclusions are drawn in Section VII.

II. CONTEST MOTIVATIONS AND SCOPE

The Data Fusion Committee (DFC) of the IEEE Geoscience and Remote Sensing Society (GRS-S) has begun sponsoring a yearly "Data Fusion Contest," in conjunction with the International Geoscience and Remote Sensing Symposium (IGARSS). Focusing every year on one specific application, the basic idea consists in distributing some data sets to the participants and specifying one specific goal. Every participating team is then allowed a few weeks to apply its algorithms to the data sets and send the corresponding results back to the organizers. All the results are then evaluated both visually and using standard metrics. A comparison that aims to note every algorithm's

specific merits presented during the data-fusion special session at the following IGARSS. Eventually, the most meriting teams/ algorithms are awarded an IEEE Certificate of Recognition.

For the 2006 contest, the focus was on pansharpening. In recent years, an ever-increasing number of pansharpening methods have been developed and published in the literature. The rising interest in this field justifies the contest organized by DFC, in order to compare recent methods on the same test set of images. The idea is to provide an overview of the performances that can be expected and help the user find the right tool. However, achieving such a comparison is quite a difficult task.

1) Objective comparison: This can only be achieved by using common data sets. Whereas authors are traditionally not allowed to further distribute the original images on which their publications are based, the organization of the contest allows the research community to share data sets, thus providing a common reference. One should note that data were not provided by any of the participant but by external institutions (see Section IV-B), thus ensuring equality: The data sets were new to every participant.

2) Representative test images: One should also note that the participants had to process whole images and not only predefined patches. As a matter of fact, this mimics real-life applications where the algorithms are designed to process one image as a whole and are not specifically tuned to match one given patch's local characteristics. Furthermore, the provided data sets included various regions (urban areas, outskirts, countryside, etc.).

3) Optimal use of each algorithm: Coding someone else's algorithm to exactly reproduce the corresponding results and test the method on other images is, in the best situations, extremely tricky and time-consuming. In most situations, even though one can only regret it, this is even simply not possible. By asking every participant to apply her/his own algorithm, the contest is solving this problem: With every method being applied by its initial conceiver, one can be sure that the potentiality of every method is optimally exploited.

4) Standard evaluation procedure: Finally, it is of the utmost importance to have a standard and recognized evaluation procedure. The procedure used in the contest is described in Section IV-A. It includes visual assessment of the results, together with a quantitative evaluation using standard numerical criteria. Providing such an evaluation and the resulting discussions is certainly the main contribution of this contest.

III. CONTEST PARTICIPANTS AND METHODS

The contest was first advertised in late 2005, and the call for participation was officially distributed in January 2006. The test images (over 4 GB of data) were made publicly available by secured file transfer protocol to the participants in February. One month later, the participants sent their results back for evaluation. Eventually, eight algorithms from seven different teams worldwide provided results in a format suitable for evaluation and comparison. Participants were aware that the evaluation would not be conducted regarding a specific classification task

but using all standard general quality measurements requiring corresponding references.

The eight algorithms are briefly described in the next sections.

A. Additive Wavelet Luminance Proportional (AWLP)

The AWLP [23] developed at the Computer Vision Center, Barcelona, Spain, relies on MRA achieved through an "à trous" wavelet transform. The algorithm is a modified version of additive wavelet (AWL) [18], which works in the IHS domain by performing MRA of the I component. To improve the spectral quality, the high-pass details are injected proportionally to the low-pass MS components in such a way that the fused MS pixel vector is always proportional to that before fusion.

B. Fast Spectral Response Function (FSRF)

The FSRF method developed at the University of Navarra, Spain, is an unpublished variant of the method in [17], which is an improved version of the fast GIHS by Tu *et al.* [16]. Substantially, FSRF is a CS method in which the spatial detail, given as difference between the Pan image and the generalized intensity, is weighted by the coefficients derived from the spectral response functions of MS and Pan.

C. GIHS With Genetic Algorithm (GIHS-GA)

The GIHS-GA [24] developed at the University of Siena, Italy, is based on CS strategy. The weights of the MS bands in synthesizing the intensity component and the injection gains are achieved by minimizing a global distortion metrics ($Q4$, in this case) by means of a GA. Minimization is carried out at a coarser scale and the model parameters (four weights and four gains for a four-band image) inferred at the finer scale. The optimization step is time-consuming and is usually performed on a statistically meaningful subimage. The same approach, but coupled with MRA, is described in [25].

D. GIHS With Tradeoff Parameter (GIHS-TP)

The GIHS-TP developed at the Advanced Institute of Science and Technology, Daejeon, Korea, is a CS-based method that trades off the performances of GIHS [16] in terms of spectral distortion and spatial enhancement [26]. By adjusting the parameter between 1 and ∞, the method yields all intermediate results between plain resampling without spatial enhancement and standard GIHS fusion.

E. Generalized Laplacian Pyramid With Context-Based Decision (GLP-CBD)

The GLP-CBD was developed at the "Nello Carrara" Institute of Applied Physics of the National Research Council, Florence, Italy. It exploits MRA, achieved through GLP [19], with the spatial frequency response of the analysis filters matching a model of the modulation transfer function (MTF) of the MS instrument [20]. The injection model employs a decision based on locally thresholding the correlation coefficient (CC)

between the resampled MS band and the low-pass approximation of the Pan.

F. University of New Brunswick (UNB)-Pansharp

The pansharpening algorithm developed at the UNB, Canada, is based on CS. The least squares technique is utilized to reduce color distortion, by identifying the best fit between gray values of individual image bands and adjusting the contribution of the individual bands to the fusion result [27]. A set of statistic approaches is employed to automate the fusion process, by estimating the gray-value relationship between all the input bands and eliminating the influence of data set variation.

G. Window Spectral Response (WiSpeR)

The WiSpeR algorithm developed at the Computer Vision Center, Barcelona, as well, is a generalization of AWLP [23]. It takes into account the relative spectral responses of MS and Pan in defining a set of weights that rule the injection of high-pass wavelet planes.

H. Weighted Sum Image Sharpening (WSIS)

The WSIS algorithm was developed by Ball Aerospace and Technologies Corporation, Fairborn, OH. The pansharpened MS bands are given as linear combination of the resampled bands and of the Pan image histogram-matched to each of the MS band. The method can be brought back to CS, even though the transformation yielding the component to substitute is unconventional.

IV. EVALUATION CRITERIA AND DATA SETS

A. Evaluation Criteria

In the framework of this contest, the performances of each method are assessed by comparison to a reference. Then, the methods are ranked according to the conclusions of the visual analysis and the results from quality budgets.

As there is no obvious reference available at high spatial resolution, original Pan and MS images were spatially degraded down to a lower resolution in order to compare the fused product to the only genuine references, formed by MS original set. This is the way to check the synthesis property [28], [29]: 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. These authors also recommend to check another property called the consistency property: It states that any synthetic image, once degraded to its original resolution, should be as close as possible to the original image. In other words, spatial degradation of the fused image should lead to the original image or close. Consistency, however, is a necessary condition, and its fulfillment does not imply a correct fusion. Many of the methods tested during this contest use multiscale approaches in order to inject high spatial frequency components while preserving low spatial frequency components. Fusion methods adopting such approaches usually check this property [2], [3]. Therefore, it was believed that

the verification of the synthesis property only would be more discriminating, and the assessment focuses on that.

Quality assessment of the pansharpened MS images is a much-debated problem [30], [31]. Even when reference MS images are available for comparisons with fusion results, assessment of fidelity to the reference usually requires computation of a number of different indexes. Examples are CC between each band of the fused and reference MS images, the bias in the mean, the root mean square error (rmse), and the average angle error, which is defined as in the following.

Given two spectral vectors $\mathbf{v}$ and $\hat{\mathbf{v}}$, both having L components, in which $\mathbf{v} = \{v_1, v_2, \ldots, v_L\}$ is the original spectral pixel vector $v_l = G^{(l)}(i, j)$, while $\hat{\mathbf{v}} = \{\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_L\}$ is the distorted vector obtained by applying fusion to the coarser resolution MS data, i.e., $\hat{v}_l = \hat{G}^{(l)}(i, j)$, the Spectral Angle Mapper (SAM) denotes the absolute value of the spectral angle between the two vectors

$$\text{SAM}(\mathbf{v}, \hat{\mathbf{v}}) \triangleq \arccos \left(\frac{\langle \mathbf{v}, \hat{\mathbf{v}} \rangle}{\|\mathbf{v}\|_2 \cdot \|\hat{\mathbf{v}}\|_2} \right). \qquad (1)$$

A value of SAM (1) equal to zero denotes absence of spectral distortion, but radiometric distortion is possible (the two pixel vectors are parallel but have different lengths). SAM is measured in either degrees or radians and is usually averaged over the whole image to yield a global measurement of spectral distortion.

In 2000, Ranchin and Wald [2] proposed an error index that offers a global picture of the quality of the fused product. This error is called relative dimensionless global error in synthesis (ERGAS) and is given by

$$\text{ERGAS} \triangleq 100 \frac{d_h}{d_l} \sqrt{\frac{1}{L} \sum_{l=1}^{L} \left(\frac{\text{rmse}(l)}{\mu(l)} \right)^2} \qquad (2)$$

where d_h/d_l is the ratio between the pixel sizes of Pan and MS, e.g., 1/4 for Ikonos and QuickBird data, $\mu(k)$ is the mean (average) of the lth band, and L is the number of bands. The ideal value of ERGAS is zero.

An image quality index suitable for MS images having four spectral bands was recently proposed in [32] to assess the pansharpening methods. The quality index $Q4$ is a generalization to four-band images of the Q index [33], which can be applied only to monochrome images. $Q4$ is obtained through the use of CC between hypercomplex numbers, or quaternions, representing spectral pixel vectors. $Q4$ is made of three different factors: The first is the modulus of the hypercomplex CC between the two spectral pixel vectors and is sensitive to both the loss of correlation and to spectral distortion between the two MS data sets. The second and third terms, respectively, measure contrast changes and mean bias on all bands simultaneously. The modulus of the hypercomplex CC measures the alignment of the spectral vectors. Therefore, its low value may detect when radiometric distortion is accompanied by spectral distortion. Thus, both radiometric and spectral distortions may be encapsulated in a unique parameter. All statistics are calculated as averages on $N \times N$ blocks, either $N = 16$ or $N = 32$. Eventually, $Q4$ is averaged over the whole image to yield the

global score index. The highest value of $Q4$, attained if and only if the test MS image is equal to the reference, is one; the lowest value is zero.

Quantitative quality assessment calls upon distances that measure the discrepancy between an image, single or multimodal, and its reference. The overall quality of the fused product is characterized by the use of a quality budget composed by a combination of several distances. Thomas and Wald [29] recently proposed a categorization of distances found in the literature based on their properties, complementarities, and redundancies. Following their recommendations, we selected a quality budget containing the following distances: relative bias in percent, relative difference of variances, relative standard deviation, and CC as unimodal statistics; ERGAS, average SAM, and average $Q4$, as multimodal indexes. Several of these indexes partly capture the resemblance of the geometrical quality because geometry is defined by differences in gray values. Several indexes were proposed in order to quantify edge sharpness, based on the point spread function [30] and on the MTF [34]. Compared to the aforementioned indexes, they are more recent and are less documented. The knowledge of their properties and their appropriate use is still poor, and we preferred not to use them. Furthermore, we consider that the visual analysis is a powerful tool in capturing this geometrical aspect, as advocated by Laporterie-Déjean *et al.* [35].

B. Data Sets

From 1999 to 2002, the Centre National d'Études Spatiales (CNES—French Space Agency) led a research program that aimed to identify and evaluate the panchromatic/MS fusion method on inframetric spatial resolution images. In this paper, images have been acquired from an airborne platform over nine sites covering various landscapes (rural, urban, or coastal). Five methods, which are selected as the most efficient, were compared in this data set. The evaluation was both qualitative with visual expert opinions and quantitative with statistical criteria [35].

The Pléiades data set was kindly made available by CNES for the contest. The four MS bands collected by the aerial platform have 60-cm resolution. Since the panchromatic camera was under development, the high-resolution panchromatic image was obtained as follows: 1) averaging the green and red channels; 2) applying the nominal MTF of the panchromatic camera; 3) resampling the outcome to 80 cm; 4) adding the instrument noise; and 5) recovering the ideal image by means of inverse filtering and wavelet denoising. The MS bands with spatial resolution of four times lower than that of the panchromatic were simulated by proper low-pass filtering and decimation. The frequency response of the low-pass filter was designed to match the MTF of each spectral channel of the spaceborne instrument. The size of the (simulated) panchromatic image is approximately 5000 × 20 000 pixels for each scene; the four reduced MS bands are 1250 × 5000. Six scenes were made available to the contest participants, who were requested to fuse all the data sets. Only two scenes, portraying the cities of Toulouse and Strasbourg, were used for evaluations, because they had been processed by (almost) all participants.

The main advantage of the simulated Pléiades data set is that fusion performances can be objectively evaluated at the same spatial scale of the final user's product, unlike in fusion assessments carried out on spatially degraded versions of commercial data. Perhaps, one disadvantage is that the simulated panchromatic image is narrowband, i.e., it does not comprise the NIR wavelengths, while the instrument which will be launched will have spectral response in 500–850 nm, analogously to other instruments (IKONOS and QuickBird). The bandwidth of the panchromatic is likely to influence the performances of fusion algorithms.

To overcome this limitation, a second data set from a commercial instrument was made available by the Mississippi State University. The MS and panchromatic image acquired by the QuickBird satellite imager on the Mississippi State University campus were downsized by four to allow quantitative evaluations to be performed. Although fusion is assessed at a scale that is different from that of the user's product (2.8 m instead of 0.7 m), the data set is still significant because the panchromatic image has not been synthesized from the MS bands, and its bandwidth also comprises the NIR wavelengths. The size of the degraded image is 1834 × 1721 (samples × lines).

V. Results

A. Evaluation Procedure

The evaluation procedure consists of the following main steps.

1) Detect two regions of interest (ROI) of size 1024 × 1024 on a subset of images comprising Pléiades—Toulouse, Pléiades—Strasbourg, and QuickBird—Mississippi State University campus (on Pléiades data, the two ROIs represent urban area and outskirts; on QuickBird data, outskirts, and countryside).
2) For each of the six patches, quality indexes (both cumulative and band-to-band) are calculated, and visual analysis is performed, compared with reference originals. Visual analysis is translated into numerical values as well.
3) A ranking of methods is produced for each patch, and such ranks are averaged over all patches and rounded to yield the overall ranking of methods.

Details of original and fused images are shown in Figs. 1 and 2 for the urban area of Toulouse and the outskirts of Mississippi State University campus, respectively.

B. Visual Analysis

The first step of quality assessment is the visual analysis of the fused images. We thus observe the main trends of errors (global scale) then analyze more precisely local artifacts. Fused color images were compared to their references. Several aspects of the image quality were taken into account. We differentiated radiometrical and geometrical viewpoints. To check the quality of the synthesized geometry, we focused on linear features, punctual objects, surfaces, edges of buildings, roads, or bridges. We analyzed the different effects named by the terms blooming, blurring, and halo that occur at global and local scales. Visual

radiometric quality regard colors of the fused color composites. We commented on the predominance of one color on another and on other spectral distortions; we specified if the phenomenon was local or global in the whole image. Other color phenomena were noticed, such as the presence of punctual colored features, saturation, or faded colors in the image. Then, the visual quality assessment is completed by observations regarding the general impression of the image: too much or not enough sharpness, good colors, or distorted ones. Then, we displayed the references and all fused products for a given image side by side in order to propose final ranking of the methods. We use previous conclusions on the radiometrical and geometrical aspects to decide between the performances of the two fused products. As the true color images did not consider the NIR band, false color representations (NIR, R, and G as R-G-B display channels) were produced, and visual inspection of this modality was performed following the same approach previously described.

As outcome of the visual analysis, an evaluation report of each patch is produced. A fused product is marked as "unacceptable" if it exhibits noticeable drawbacks affecting the whole image, such as global spectral distortions, pronounced halo, numerous noisy areas. The product is marked as "unacceptable for detailed analysis" if it does not show the above drawbacks but displays small-scale artifacts impeding reliable analysis of objects. The visual evaluation phase is summarized in the following list, where the judgements on each method are merged and synthesized.

1) *AWLP:* Pretty nice as a whole. Efforts to do with colors. Slightly less good than GLP-CBD. Small size objects are missing, particularly those in red or blue. Bias and large amount of variance introduced are observed for Strasbourg. Unacceptable for detailed visual analysis.
2) *FSRF:* Contours are not sharp enough. Colors are not synthesized well enough. Small objects, such as cars, are white and not colored. Other colors are missing in detail, thus preventing an accurate reading of small objects. Bias is observed at times. Unacceptable for detailed visual analysis.
3) *GIHS-GA:* As a whole, colors are not correctly synthesized enough. Small colored objects are missing, particularly in blue and red. Contours are blurred and not sharp enough. Bias is observed for Strasbourg but not for QB. Large amount of variance is missing. Unacceptable for detailed visual analysis.
4) *GIHS-TP:* Image is blurred but with too enhanced contours at times. Colors are respected as a whole but not for small objects. Shapes are very blurred, even for large objects. Saturation occurs, i.e., objects are often too white. Highly contrasted objects are surrounded by halo. Bias is observed. No enough variance is introduced for all bands. There is large standard deviation at pixel level. Unacceptable.
5) *GLP-CBD:* Image is nice as a whole. Colors should be better synthesized. This would enhance the legibility of the image. Details are there, except for the most colored
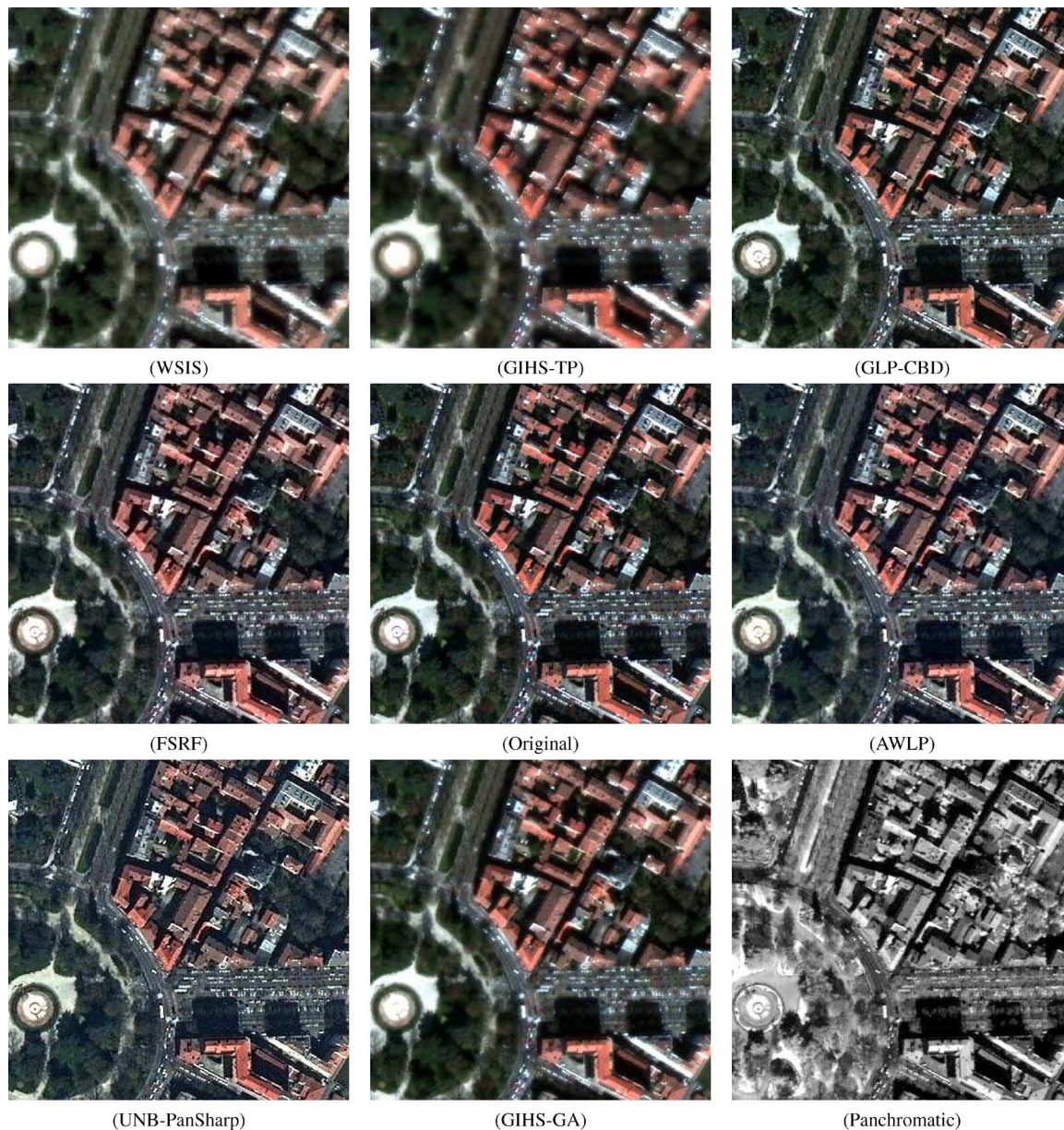
Fig. 1. Results of the fusion algorithms displayed as 352×352 true color compositions at 0.8-m pixel spacing for the simulated Pléiades image of Toulouse, urban ROI. The reference 0.8-m original MS bands and the Pan image are also displayed.

(blue, red). Errors in colors lead to interpretation errors. Contours should be sharper. There is no bias, except for Strasbourg outskirts. Unacceptable for detailed visual analysis.

6) *UNB-Pansharp:* Image is too noisy. There are many artifacts. Colors are not well synthesized as a whole and locally. Green trees are not green enough. Red or blue cars are absent. Shapes are not well defined; they are sometimes underlined by black lines. Too large bias is observed. There is lack of variance as a whole. At times, unacceptable. In best cases, unacceptable for detailed visual analysis.

7) *WiSpeR:* Image is too bluish. In one case, green is missing; in the other case, it is red. Contours are artificially enhanced. There are many artifacts due to colors. Missing colors lead to errors in interpretation; colored objects are missing. There is no bias. There is too much variance introduced. Unacceptable.

8) *WSIS:* Image is often blurred. Sometimes noisy (QB Countryside). Contours are not sharp enough. At times, objects are saturated. Colors are not well synthesized. A large bias is observed, which means there is a discrepancy in spectral content. Large amount of variance is missing. Unacceptable. In best case (QB Countryside), unacceptable for detailed visual analysis.

C. Quantitative Evaluation

All statistical indexes, both unimodal (percentage bias, relative difference of variances, relative standard deviation of
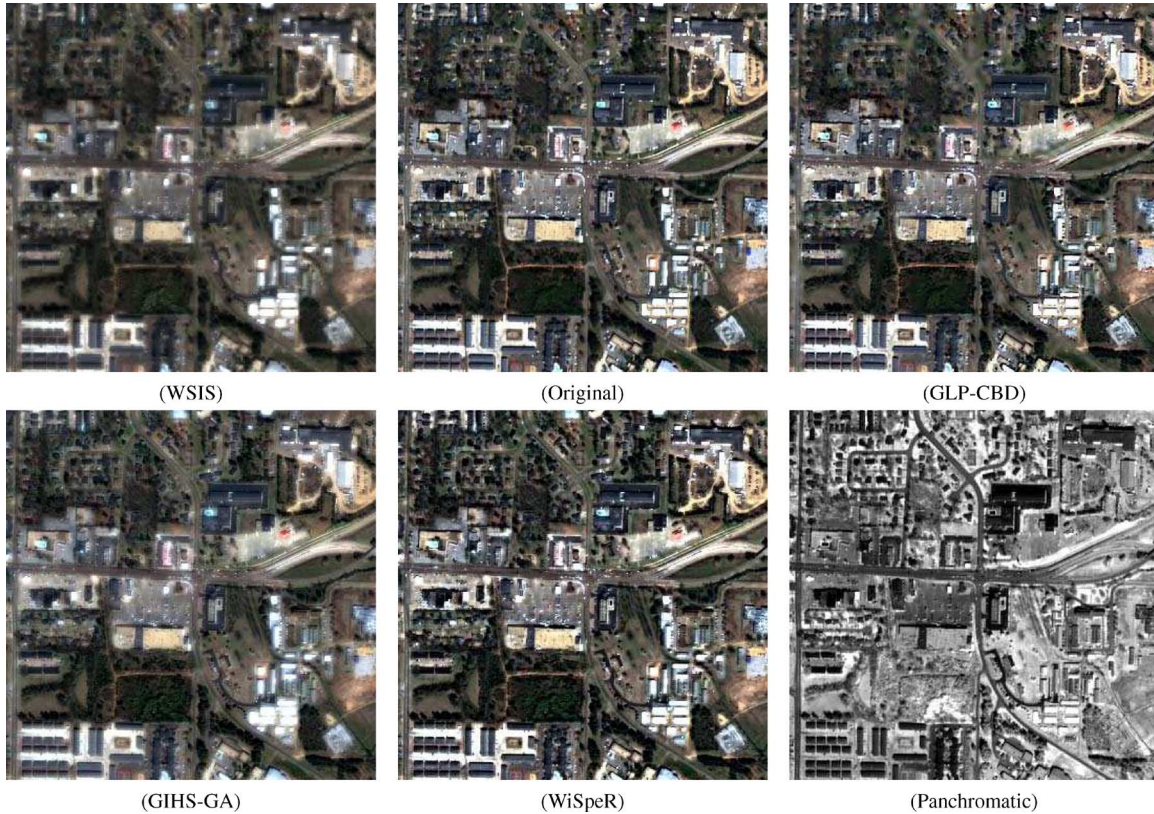
Fig. 2.   Results of the fusion algorithms displayed as 352×352 true color compositions at 2.8-m pixel spacing for the degraded QuickBird image of Mississippi State University campus, outskirts ROI. The reference 2.8-m original MS bands and the degraded 2.8-m Pan image are also displayed.

TABLE I
TEST-IMAGE TOULOUSE URBAN: AVERAGE CUMULATIVE QUALITY/DISTORTION INDEXES BETWEEN ORIGINAL 0.8-m MS BANDS AND FUSED IMAGES OBTAINED FROM 3.2-m MS AND 0.8-m PAN

	AWLP	FSRF	GIHS-GA	GIHS-TP	GLP-CBD	UNB-Pansharp	WiSpeR	WSIS
Q4	0.96	0.95	0.93	0.84	0.96	0.90	-	0.86
ERGAS	3.12	3.50	3.63	5.34	2.78	4.68	-	5.59
SAM(°)	4.56	5.74	4.08	5.30	3.67	5.40	-	6.05

TABLE II
TEST-IMAGE QUICKBIRD OUTSKIRTS: AVERAGE CUMULATIVE QUALITY/DISTORTION INDEXES BETWEEN ORIGINAL 2.8-m MS BANDS AND FUSED IMAGES OBTAINED FROM 11.2-m MS AND 2.8-m PAN

	AWLP	FSRF	GIHS-GA	GIHS-TP	GLP-CBD	UNB-Pansharp	WiSpeR	WSIS
Q4	-	-	0.90	-	0.94	-	0.90	0.84
ERGAS	-	-	2.83	-	2.13	-	3.10	3.62
SAM(°)	-	-	3.95	-	2.98	-	3.34	4.60

the difference image, CC) and multimodal ($Q4$, ERGAS, and average SAM), have been calculated for all the six test patches. Results of the multimodal evaluation are reported in Tables I and II, and are relative to the test images shown in Figs. 1 and 2, respectively.

As shown in Tables I and II, some participants have chosen not to apply a method to all data sets or to apply different methods to the data that originated from different instruments. In some cases, the reason may have been computational complexity, because participants were requested to process large images in a short time. In another case of different methods

from the same team, it is also possible that none of the methods were adequately performing on both data sets.

The ranking of methods per image patch, obtained by combining unimodal scores, multimodal scores, and visual judgements into a unique score and sorting the results by decreasing score, would be the one reported in Table III. Clearly, it impossible to get the final rank of methods simply by averaging the partial ranks achieved on individual patches. In order to make a homogeneous, yet unbiased, comparisons, the missing numerical values, e.g., those in Tables I and II, have been filled with the average of those obtained by the other methods

TABLE III
RANKING OF METHODS PER TEST IMAGE EACH RANK IS RELATIVE TO THE NUMBER OF ALGORITHMS THAT HAVE PROCESSED THAT PATCH

	AWLP	FSRF	GIHS-GA	GIHS-TP	GLP-CBD	UNB-Pansharp	WiSpeR	WSIS
Toulouse Outskirts	2/7	3/7	4/7	7/7	1/7	5/7	-	6/7
Toulouse Urban	2/7	3/7	4/7	7/7	1/7	5/7	-	6/7
Strasbourg Outskirts	2/6	4/6	3/6	6/6	1/6	5/6	-	-
Strasbourg Urban	2/6	4/6	3/6	6/6	1/6	5/6	-	-
Quickbird Outskirts	-	-	2/4	-	1/4	-	3/4	4/4
Quickbird Countryside	-	-	1/4	-	2/4	-	4/4	2/4

available on the same patch. After that, scores relative to individual measurements are synthesized into scores cumulative of all indexes and translated into ranks; eventually, ranks are averaged over all the six patches to yield the overall rank of each of the eight methods.

D. Ranking of Methods

The final ranking of the methods is as follows:

1) GLP-CBD;
2) AWLP;
3) GIHS-GA;
4) WiSpeR;
5) FSRF;
6) UNB-Pansharp;
7) WSIS;
8) GIHS-TP.

We would like to point out that the eight methods are not uniformly spaced in performances. Four classes can be recognized, such that performances of the methods belonging to the same class are more similar than those of the methods belonging to neighboring classes. The first class comprises GLP-CBD and AWLP. These two algorithms, clearly outperforming all the others, received an IEEE Certificate of Recognition during IGARSS'06 in Denver. The second class includes GIHS-GA, WiSpeR, and FSRF. UNB-Pansharp and WSIS make up the third class, while GIHS-TP is an outlier.

The WSIS method performs better on countryside regions and can be of interest depending on applications. The UNB-Pansharp method was specifically designed for broadband Pan (like Ikonos and QuickBird) and is likely to have been penalized in the evaluations because the degraded QuickBird image has not been fused by this method.

Evaluation results highlight that MRA-based methods (GLP-CBD, AWLP, and WiSpeR) are on an average more performing than CS-based methods (GIHS-GA, FSRF, UNB-Pansharp, WSIS, and GIHS-TP), which, on the other hand, have the advantage of being simpler and suitable for fast implementations. An exception in this sense is GIHS-GA, which, however, is the most performing CS method among those compared.

VI. DISCUSSION

The outstanding performances of GLP-CBD and AWLP can be explained by the fact that the two methods rely on very similar concepts and employ similar tools to achieve the goal of pansharpening.

Both methods exploit MRA to achieve spatial enhancement. MRA is provided by the GLP for the former and by the "à trous" wavelet transform for the latter. For a GLP with expansion filter that is almost ideal, as it is [19], the frequency responses of the equivalent analysis filter banks of GLP and "à trous" wavelet, whose high-pass filter selects details that are injected, depend only on the analysis filter. If such a filter is designed in such a way that it matches the MTF of the sensor acquiring the band that is being enhanced, the complementary high-pass filter will inject the exact amount of frequency components that have not been selected by the optical filter of the instrument. MRA is adjusted to the MTFs of the MS scanner (one for each band), either directly, as for GLP-CBD [20], or indirectly, as for AWLP, which utilizes Starck and Murtagh Gaussian-like cubic spline filter [18], whose frequency response matches the typical shape of an isotropic MTF, with an amplitude that is approximately equal to 0.2 at the cutoff Nyquist frequency.

The detail injection models are different but are both adaptive, i.e., data-dependent. The former (CBD) aims to locally unmix the coarse MS pixels through the sharp Pan image by means of a space-varying injection gain comprising the local CC between resampled MS and low-pass filtered Pan. The latter aims to minimize the spectral distortion by injecting a detail vector parallel to the resampled MS vector at each pixel. In this way, colors are preserved in terms of hue, while saturation may be changed.

VII. CONCLUSION

In this paper, we presented the outcome of the 2006 issue of the Data Fusion contest organized by the IEEE GRS-S DFC. The focus of this first issue was on pansharpening algorithms, i.e., fusion of images with different spatial resolutions in the special case of optical satellite imagery. The goal was to increase the spatial resolution of MS images up to the high resolution of the available panchromatic image while preserving the original spectral resolution. With a standardized evaluation procedure, including visual and quantitative analyses, the results of the eight algorithms provided by the seven different research groups have been compared. The participants blindly applied their algorithms to complete the common images from different sensors (namely QuickBird and simulated Pléiades), featuring different regions (urban areas, outskirts, countryside), without knowing any predefined patch.

The visual analysis and the extensive numerical comparisons (over 600 values of statistical indexes have been computed and synthesized) consistently concluded to the superiority of the two algorithms that outperform all the others. It is interesting to note that the algorithms, based on MRA, generally performed

better than the ones based on CS. Furthermore, the two best algorithms share a similar philosophy, taking into account the instrument-related physical models (MTF). This surely points to a way for future developments of pansharpening algorithms.

ACKNOWLEDGMENT

REFERENCES

[1] L. Wald, "Some terms of reference in data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1190–1193, May 1999.

[2] T. Ranchin and L. Wald, "Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation," *Photogramm. Eng. Remote Sens.*, vol. 66, no. 1, pp. 49–61, Jan. 2000.

[3] 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.

[4] I. Couloigner, T. Ranchin, V. P. Valtonen, and L. Wald, "Benefit of the future SPOT-5 and of data fusion to urban road mapping," *Int. J. Remote Sens.*, vol. 19, no. 8, pp. 1519–1532, 1998.

[5] V. S. Raptis, R. A. Vaughan, T. Ranchin, and L. Wald, "Assessment of different data fusion methods for the classification of an urban environment," in *Proc. 2nd Conf. "Fusion of Earth Data: Merging Point Measurements, Raster Maps Remotely Sensed Images,"* T. Ranchin and L. Wald, Eds., 1998, pp. 167–182.

[6] L. Wald and J.-M. Baleynaud, "Observing air quality over the city of Nantes by means of Landsat thermal infrared data," *Int. J. Remote Sens.*, vol. 20, no. 35, pp. 947–959, 1999.

[7] W. Yang, F. Cauneau, J.-P. Paris, and T. Ranchin, "Influence of landscape changes on the results of the fusion of P and XS images by different methods," in *Proc. 3rd Conf. "Fusion of Earth Data: Merging Point Measurements, Raster Maps Remotely Sensed Images,"* T. Ranchin and L. Wald, Eds., 2000, pp. 47–56.

[8] Y. Cornet, S. de Béthune, M. Binard, F. Muller, G. Legros, and I. Nadasdi, "RS data fusion by local mean and variance matching algorithms: Their respective efficiency in a complex urban context," in *Proc. Joint IEEE/ISPRS Workshop Remote Sens. Data Fusion Over Urban Areas*, 2001, pp. 105–111.

[9] A. Fanelli, A. Leo, and M. Ferri, "Remote sensing images data fusion: A wavelet transform approach for urban analysis," in *Proc. Joint IEEE/ISPRS Workshop Remote Sens. Data Fusion Over Urban Areas*, 2001, pp. 112–116.

[10] D. Vaiopoulos, K. Nikolakopoulos, and G. Skianis, "A comparative study of resolution merge techniques and their efficiency in processing image of urban areas," in *Proc. Joint IEEE/ISPRS Workshop Remote Sens. Data Fusion Over Urban Areas*, 2001, pp. 270–274.

[11] M. Gonzáles Audícana and A. Seco, "Fusion of multispectral and panchromatic images using wavelet transform. Evaluation of crop classification accuracy," in *Proc. 22nd EARSeL Annu. Symp. Geoinformation Eur.-Wide Integr.*, Prague, Czech Republic, 4–6 June 2002, T. Benes, Ed., 2003, pp. 265–272.

[12] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, "A comparative analysis of image fusion methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1391–1402, Jun. 2005.

[13] V. K. Shettigara, "A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set," *Photogramm. Eng. Remote Sens.*, vol. 58, no. 5, pp. 561–567, May 1992.

[14] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, "A new look at IHS-like image fusion methods," *Inf. Fusion*, vol. 2, no. 3, pp. 177–186, Sep. 2001.

[15] Y. Zhang, "Understanding image fusion," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 6, pp. 657–661, Jun. 2004.

[16] 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. Remote Sens. Lett.*, vol. 1, no. 4, pp. 309–312, Oct. 2004.

[17] M. Gonzáles Audícana, X. Otazu, O. Fors, and J. A. Alvarez-Mozos, "A low computational-cost method to fuse IKONOS images using the spectral response function of its sensors," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1683–1691, Jun. 2006.

[18] J. Núñez, X. Otazu, O. Fors, A. Prades, V. Palà, and R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1204–1211, May 1999.

[19] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-driven fusion of high spatial and spectral resolution data based on oversampled multiresolution analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2300–2312, Oct. 2002.

[20] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-tailored multiscale fusion of high-resolution MS and Pan imagery," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, May 2006.

[21] M. Gonzáles Audícana, J. L. Saleta, R. García Catalán, and R. García, "Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1291–1299, Jun. 2004.

[22] A. Garzelli and F. Nencini, "Interband structure modeling for Pan-sharpening of very high resolution multispectral images," *Inf. Fusion*, vol. 6, no. 3, pp. 213–224, Sep. 2005.

[23] X. Otazu, M. Gonzáles Audícana, O. Fors, and J. Núñez, "Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2376–2385, Oct. 2005.

[24] A. Garzelli and F. Nencini, "Fusion of panchromatic and multispectral images by genetic algorithms," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2006, pp. 3810–3813.

[25] A. Garzelli and F. Nencini, "PAN-sharpening of very high resolution multispectral images using genetic algorithms," *Int. J. Remote Sens.*, vol. 27, no. 15, pp. 3273–3292, Aug. 10, 2006.

[26] M. Choi, "A new intensity–hue–saturation fusion approach to image fusion with a tradeoff parameter," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1672–1682, Jun. 2006.

[27] Y. Zhang, "A new automatic approach for effectively fusing Landsat-7 as well as IKONOS images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2002, pp. 2429–2431.

[28] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogramm. Eng. Remote Sens.*, vol. 63, no. 6, pp. 691–699, 1997.

[29] C. Thomas and L. Wald, "Comparing distances for quality assessment of fused products," in *Proc. 26th EARSeL Annu. Symp. New Develop. Challenges Remote Sens.,* Warsaw, Poland, May 29–31, 2006. Z. Bochenek, Ed., Rotterdam, The Netherlands: Balkema, 2007, pp. 101–111.

[30] J. Li, "Spatial quality evaluation of fusion of different resolution images," *ISPRS Int. Arch. Photogramm. Remote Sens.*, vol. 33, no. B2-2, pp. 339–346, 2000.

[31] C. Thomas and L. Wald, "Assessment of the quality of fused products," in *Proc. 24th EARSeL Annu. Symp. New Strategies Eur. Remote Sens.,* Dubrovnik, Croatia, May 25–27, 2004. M. Oluic, Ed., Rotterdam, The Netherlands: Balkema, 2005, pp. 317–325.

[32] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of Pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, Oct. 2004.

[33] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[34] C. Thomas and L. Wald, "An MTF-based distance for the assessment of the geometrical quality of fused products," in *Proc. 9th Int. Conf. Inf. Fusion*, Florence, Italy, Jul. 10–13, 2006, pp. 1–7.

[35] F. Laporterie-Déjean, H. de Boissezon, G. Flouzat, and M.-J. Lefèvre-Fonollosa, "Thematic and statistical evaluations of five panchromatic/multispectral fusion methods on simulated PLEIADES-HR images," *Inf. Fusion*, vol. 6, no. 3, pp. 193–212, Sep. 2005.

Luciano Alparone received the Laurea degree (*cum laude*) in electronic engineering and the Ph.D. degree from the University of Florence, Firenze, Italy, in 1985 and 1990, respectively.

Since 2002, he has been an Associate Professor of electrical communications with the Dipartimento di Elettronica e Telecomunicazioni, Università di Firenze, Firenze. During the spring of 2000 and summer of 2001, he was a Visiting Researcher with the Tampere International Centre for Signal Processing, Tampere, Finland. His research interests are data compression for remote sensing and medical applications, multiresolution image analysis and processing, multisensor data fusion, and processing and analysis of SAR images. He authored or coauthored over 60 papers on peer-reviewed journals and a total of almost 300 publications.

Dr. Alparone was a recipient of the 2004 Geoscience and Remote Sensing Letters Prize Paper Award for his study on "A global quality measurement of Pan-sharpened multispectral imagery."

Lucien Wald received the degree in theoretical physics from the University of Marseille, Marseille, Paris, and the University of Paris, Paris, France, in 1977, the Ph.D. degree in applications of remote sensing to oceanography from the University of Paris in 1980, and the Doctorat d'Etat ès Sciences degree in applications of remote sensing to oceanography from the University of Toulon, Toulon, France, in 1985.

Since 1991, he has been a Professor with the Groupe Teledetection and Modelisation, Centre d'Energetique, Ecole des Mines de Paris, Sophia Antipolis, France. His research interests include applied mathematics and meteorology.

Dr. Wald received the Autometrics Award in 1998 and the Erdas Award in 2001 for his articles on data fusion. His career in information technologies was awarded in 1996 by the famous French Blondel Medal.
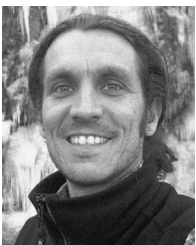
Jocelyn Chanussot (M'04–SM'04) received the degree in electrical engineering from the Grenoble Institute of Technology (INP Grenoble), Grenoble, France, in 1995, and the Ph.D. degree from the University of Savoie, Annecy, France, in 1998.

In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement (French National Defense Department). Since 1999, he has been teaching signal and image processing with the INP Grenoble and has been working at the Grenoble Images sPeech Signals and Automatics Laboratory (GIPSA-Lab), Grenoble, as an Assistant in 1999–2005, an Associate in 2005–2007, and currently a Full Professor. His research interests include statistical modeling, multicomponent image processing, nonlinear filtering, remote sensing, and data fusion. He is an Associate Editor for the journal *Pattern Recognition* (2006–2008).

Dr. Chanussot has been an Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (2005–2007) and is currently an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (2007–2010). He is the Cochair of the IEEE Geoscience and Remote Sensing Society Data Fusion Technical Committee (2005–2007) and a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Beta Society (2006–2008).

Claire Thomas received the degree in physics from the Ecole Nationale Supérieure de Physique de Strasbourg, Strasbourg, France, in 2002, the Master's degree in photonics, image, and cybernetics from the University of Strasbourg, Strasbourg, in 2002, and the Ph.D. degree from Ecole des Mines de Paris, Sophia Antipolis, France, in 2006, focusing on the fusion of images of different spatial resolutions.

She was with the National Severe Storm Laboratory, Norman, OK, for a six-month internship. She is currently with the Groupe Teledetection and Modelisation, Centre d'Energetique, Ecole des Mines de Paris. She has presented several communications and published two peer-reviewed articles.

Paolo Gamba (S'91–M'93–SM'00) received the Laurea (*cum laude*) and Ph.D. degrees in electronic engineering from the University of Pavia, Pavia, Italy, in 1989 and 1993, respectively.

He is currently an Associate Professor of telecommunications with the Dipartimento di Elettronica, Università di Pavia. He has published more than 40 papers in peer-review journals on urban remote sensing and presented more than 100 papers at workshops and conferences. He has been a Guest Editor of special issues of *ISPRS Journal of Photogrammetry and Remote Sensing*, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, *International Journal of Information Fusion*, and *Pattern Recognition Letters*.

Dr. Gamba is the Organizer and Technical Chair of the biennial Geoscience and Remote Sensing Society (GRSS)/International Society for Photogrammetry and Remote Sensing (ISPRS) Joint Workshops on "Remote Sensing and Data Fusion over Urban Areas" from 2001 to 2007. He has been the Chair of Technical Committee 7 "Pattern Recognition in Remote Sensing" of the International Association for Pattern Recognition from October 2002 to October 2004. He serves as an Associate Editor of IEEE GEOSCIENCE AND REMOTE SENSING LETTERS and is currently the Chair of the Data Fusion Committee of the IEEE GRSS.

Lori Mann Bruce (S'91–M'96–SM'01) received the B.S.E. degree from the University of Alabama, Huntsville, in 1991, the M.S. degree from the Georgia Institute of Technology, Atlanta, in 1992, and the Ph.D. degree from the University of Alabama, Huntsville, in 1996, all in electrical engineering.

Currently, she is a Professor with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, where she also serves as Associate Director for Research in the Geo-Resources Institute. Her research interests include applying advanced digital signal processing techniques, such as discrete wavelet transforms, to automated pattern recognition in hyperspectral remote sensing and digital mammography.

Dr. Bruce is a member of Eta Kappa Nu, Phi Kappa Phi, and Tau Beta Pi.