

A new pan-sharpening method using multiobjective particle swarm optimization and the shiftable contourlet transform

Jamal Saeedi, Karim Faez *

Electrical Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), 424, Hafez Ave., Tehran, Iran

ARTICLE INFO

Article history:

Received 14 July 2010

Received in revised form

16 January 2011

Accepted 17 January 2011

Available online 11 February 2011

Keywords:

Pan-sharpening

Shiftable contourlet transform

Multiobjective particle swarm optimization

ABSTRACT

In this paper, a novel approach based on multiobjective particle swarm optimization (MOPSO) is presented for panchromatic (Pan) sharpening of a multispectral (MS) image. This new method could transfer spatial details of the Pan image into a high-resolution version of the MS image, while color information from the low-resolution MS image is well preserved. The Pan and MS images are locally different because of different resolutions, and therefore we cannot directly combine them in the spatial domain. For this reason, we generate two initial results, which are more appropriate for a weighted combination. First, the Pan and the MS images are histogram matched. Then we use the shiftable contourlet transform (SCT) to decompose the histogram-matched Pan and MS images. The SCT is a new shiftable and modified version of the contourlet transform. In this step, an algorithm based on the SCT is used to generate two initial results of the high-resolution MS images. Our objective is to produce two modified high-resolution MS images, in which one has high spatial similarity to the Pan image and the other one has high radiometric quality in each band. Therefore, we have used two different fusion rules to integrate the high-frequency contourlet coefficients of the Pan and MS images to generate two initial results of high-resolution MS image or the Pan-sharpened (PS) image. Finally, we can find the optimal PS image by applying the MOPSO algorithm and using the two initial PS results. Specifically, the PS image is obtained via a weighted combination of the two initial results, in which the weights are locally estimated via a multiobjective particle swarm optimization algorithm to generate a PS image with high spatial and radiometric qualities. Based on experimental results obtained, the produced Pan-sharpened image also has good spectral quality. The efficiency of the proposed method is tested by performing Pan-sharpening of high-resolution (Quickbird and Worldview2) and medium-resolution (Landsat-7 ETM+) datasets. Extensive comparisons with the state-of-the-art Pan-sharpening algorithms indicate that our new method provides improved subjective and objective results.

© 2011 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

1. Introduction

Pan-sharpening is a shorthand for panchromatic (Pan) sharpening. It indicates using a Pan image to sharpen a multispectral (MS) image. In this sense, to sharpen means to increase the spatial resolution of an MS image. Therefore, the goal of Pan-sharpening is to combine the high spatial resolution of the Pan image with the precise spectral information of the MS image.

An MS image is an image which contains more than one spectral band. A color image is a very simple example of an MS image which contains three bands. In this case, these bands correspond to the blue, green, and red wavelengths of the electromagnetic spectrum. In addition, many satellite imaging systems, such as the Quickbird and Landsat-7 satellites, produce a Pan image to accompany the MS

imagery. This Pan image has higher spatial resolution than the MS data, but the spectral response spans a wide range, such as visible or thermal infrared.

Pan-sharpening algorithms depend on the input images being co-registered, because they all perform operations on corresponding pixels in both images. They all do something with the MS pixel and the Pan pixels to create new pixels. If the images are not co-registered, the processing will use the wrong pixels, not the corresponding ones, and the result will not look natural. In addition, the MS data is resampled into the same spatial reference and grid as the Pan data, using nearest neighbor, bilinear, or cubic convolution techniques.

A ground-truth usually is not available to evaluate the efficiency of a Pan-sharpening technique. Therefore, quality assessment should refer to the goals of the fusion process (Wald, 1999). The most straightforward objective is image analysis, but also such automated tasks as feature extraction and segmentation or classification have been found, which have benefited from

* Corresponding author. Tel.: +98 935 3592459; fax: +98 21 6640 6469.

E-mail addresses: jamal.saeedi@yahoo.com (J. Saeedi), kfaez@aut.ac.ir (K. Faez).

pan-sharpening methods (Wahlen, 2002; Zhang and Wang, 2004; Colditz et al., 2006). Here, we categorize pan-sharpening algorithms into the four groups.

(1) Color transformation.

IHS transform: IHS (intensity–hue–saturation) is the most common image fusion method for remote sensing applications, and it is used in commercial pan-sharpening software. This method converts a color MS image from RGB space to the IHS color space. Here, the I (intensity) band is replaced by the pan image. Before fusing the images, the MS image and the pan image are histogram matched (Choi, 2006).

(2) Statistical methods.

Principal component analysis (PCA): In PCA-based methods, the PCA transform converts intercorrelated MS bands into a new set of uncorrelated components. It is assumed that the first PC image with the highest variance contains the highest amount of information from the original image and will be the ideal choice to replace the high spatial resolution pan image. All the other new MS bands are unchanged. An inverse PCA transform is performed on the modified pan and MS images to obtain a high-resolution pan-sharpened image (Vijay et al., 2008).

(3) Numerical methods.

Brovey: The basic process of the Brovey transform is as follows. First, it multiplies each MS band by the high-resolution pan band, and then divides each product by the sum of the MS bands (Pohl and Genderen, 1998).

Wavelet-based sharpening: In wavelet-based schemes, detailed information is extracted from the pan image using wavelet transforms, and is injected into the MS image. The low-frequency subband of the pan-sharpened image is usually selected from the low-resolution MS image. There are three ways to fuse high-frequency (HF) subbands: substitutive, additive, and weighted methods (Amolins et al., 2007). In the substitutive scheme, the HF subbands of the MS image are replaced with HF subbands of the pan image. In the additive scheme, the HF subbands of the pan image are added to HF subbands of the MS image, and finally, in the weighted scheme, the HF subband of the pan-sharpened image is obtained by a weighted combination of HF subbands from the pan and MS images (Ranchin and Wald, 2000). In addition, other types of multiresolution analysis have been applied instead of a wavelet transform, such as the dual-tree discrete wavelet transform (Ioannidou and Karathanassi, 2007), Laplacian pyramid (Aiazzi et al., 2002), and contourlet transform (Vijay et al., 2008).

P + XS: P + XS is a variational method, which calculates the pan-sharpened image by minimizing an energy function. It obtains the edge information of the pan image, using a gradient operator. The spectral information is obtained by approximating the pan image as a linear combination of the MS bands (Ballester et al., 2006).

(4) Hybrid methods.

Ehlers Fusion: Ehlers fusion is based on an IHS transform joined with a Fourier domain filtering. First, the IHS transform is applied to the MS bands. Using fast Fourier transform (FFT) methods, the spatial components to be enhanced can be obtained directly. The intensity spectrum is filtered with a low-pass filter whereas the pan spectrum is filtered with an inverse high-pass filter. After filtering, the images are transformed back into the spatial domain with an inverse FFT and added together to form a fused intensity component with the low-frequency information from the low-resolution MS image and the high-frequency information from the pan image. This new intensity component and the original hue and saturation components of the MS image form a new IHS image. Finally, the pan-sharpened image is obtained via an inverse IHS transform (Ehlers, 2004).

IHS + wavelet-based sharpening: Similarly to the Ehlers method, in the first step, the IHS transform is applied to extract an intensity image from the MS bands. Then the wavelet-based sharpening method is used to extract detailed information from the pan image to inject into the intensity image. This new intensity component and the original hue and saturation components of the MS image form a new IHS image. Finally, the pan-sharpened image is obtained via an inverse IHS transform (Audicana et al., 2004).

Many research papers have reported the limitations of existing fusion methods. The most significant problem is color distortion. Since 1998, further scientific papers on image fusion have been published with the emphasis on improving the fusion quality and reducing color distortion or increasing the spectral and radiometric qualities (Zhang, 2004).

In this paper, we propose a new hybrid method, in which the objective is to maintain the spatial similarity of the pan-sharpened image to the pan image and radiometric quality in each band. The pan and MS images are locally different because of different resolutions, and therefore we cannot directly combine them in the spatial domain. For this reason, we have proposed a two-step algorithm, in which two initial results of the high-resolution MS images are generated in the first step. Then, in the second step, the final pan-sharpened image is obtained through a weighted combination of the two initial results using an optimization algorithm. Here, we have used a multiresolution-based algorithm to generate the two initial results of the high-resolution MS or pan-sharpened images. The objective is to produce two modified high-resolution MS images, in which one has high spatial similarity to the pan image and the other has high radiometric quality. For this purpose, we have used the shiftable contourlet transform (SCT), which is a new shiftable and modified version of the contourlet transform, proposed by Nguyen and Oraintara (2008).

There are several ways of extraction of the spatial details of a pan image. In the literature, the discrete wavelet transform (DWT) is a common technique for this purpose (e.g. Amolins et al., 2007; Ranchin and Wald, 2000). The DWT provides a good time-frequency analysis of the signal, with a non-redundant signal representation and an optimal representation of singularities. However, the DWT suffers from five fundamental shortcomings (Selesnick et al., 2005): oscillations, aliasing, shift variance, poor directionality, and absence of phase information. Shift invariance and directional selectivity are essential to the quality of wavelet-based image fusion. Because of the downsampling operation in the DWT filter banks (FBs), the DWT is not shift invariant, and it causes some visual artifacts in the pan-sharpened image. The human visual system is primarily sensitive to moving light stimulus; therefore, moving artifacts are highly distracting to a human observer (Rockinger, 1997). In addition, if the directional selectivity of an FB is defined as the ability to extract directional features into separate images, a two-dimensional (2D) DWT has very poor directional selectivity, because a 2D DWT has four subimages, which are usually referred to as LL, LH, HL, and HH images.

The contourlet transform (CT) is an alternative multiresolution method, which provides an efficient directional representation and is efficient in capturing intrinsic geometrical structures of the natural image along the smooth contours (Do and Vetterli, 2005). Similar to the DWT, the CT is not shift invariant, and it results in artifacts along the edges to some extent. The non-subsampled contourlet transform (NSCT) was proposed to meet the shift invariance. The NSCT is the shift-invariant version of the CT, and it is built based on non-subsampled pyramid decomposition and non-subsampled filter banks, to obtain the shift invariance (Da Cunha et al., 2006; Hui and Cheng, 2008). However, the NSCT has the overcomplete ratio of $N \times K$, where N is the number of resolution levels and K is the number of directions, which is very high compared to the overcomplete ratio of the CT

(4/3). However, a reduced form of translation invariance exists, namely, energy shift invariance or “shiftability” (Simoncelli et al., 1992), which means that the energy of the output signal is shift invariant. For these reasons, we use the SCT, which is shown to have a number of desirable properties for image analysis including shiftable subband, arbitrarily high directionality, and low redundancy (11/3) (Nguyen and Oraintara, 2008).

The SCT is applied for decomposing the histogram-matched pan and MS images into different resolutions and directions. In this step, two initial pan-sharpened images are generated using the maximum absolute selection and new sign-based averaging fusion rules. Finally, the pan-sharpened image is obtained via a weighted combination of the two initial results, in which the weights are locally estimated via a multiobjective particle swarm optimization algorithm to generate a pan-sharpened image with high spatial and radiometric qualities. In the literature, the optimization problem is used for pan-sharpening, and the objective index is usually the mean square error (MSE). However, having only one objective is too simple to meet the real demands (e.g. Garzelli et al., 2008; Garzelli and Nencini, 2006). In fact, there are various kinds of evaluation index, and different indexes may be compatible or incompatible with one another, so a good evaluation index system of pan-sharpening must balance the advantages of different indexes. For this reason, we use the MOPSO algorithm for the optimization problem (Li, 2003; Coello et al., 2004). The MOPSO algorithm is adopted to generate a pan-sharpened image with high spatial and radiometric qualities based on global image fusion metrics. In this study, we have used image data from different satellite imagery consisting of Landsat-7 ETM + (28.5 m MS; 14.25 m pan), Quickbird (2.44 m MS; 60 cm pan), and Wordview2 (1.84 m MS; 46 cm pan).

This paper is structured as follows. In Section 2, we briefly describe the shiftable contourlet transform. In Section 3, the proposed pan-sharpening algorithm is presented. Section 4 gives various results and the comparisons. Finally, we conclude with a brief summary in Section 5.

2. The shiftable contourlet transform

The idea of combining the directional filter bank (DFB) and a multiresolution decomposition to create a multiresolution multidirectional decomposition was proposed by Do and Vetterli (2001). In particular, Do and Vetterli (2005) built the theory of the contourlet transform from the pyramidal DFB. The contourlet provides a new system representation for image analysis. The name contourlet was given because of its ability to capture and link the points of discontinuities to form a linear structure (contours). The two-stage process used to derive the contourlet coefficients involves a multiscale transform and a local directional transform. The points of discontinuities and multiscale transformation are obtained via a Laplacian pyramid. The local directional filter bank is used to group these wavelet-like coefficients to obtain a smooth contour. Contourlet provide $2l$ directions at each scale, where l is the number of required orientations. This flexibility of having different numbers of directions at each scale makes contourlet different from other available multiscale and directional image representations, including 2D Gabor wavelets (Daugman, 1980), cortex transforms (Watson, 1987), and the steerable pyramid (Simoncelli et al., 1992).

As mentioned above, the CT is shift variant, and therefore we have used a shift-invariant version of it in this paper. The shiftable contourlet transform or shiftable complex directional pyramid is a novel image transform, which provides a directional and shiftable image representation, as shown by its ideal construction in Fig. 1. The forward and inverse transforms are carried out by the analysis and synthesis pyramidal dual-tree directional filter bank (PDTDFB).

The PDTDFB is a combination of an iterative multiresolution FB and a dual-tree DFB at each resolution, as presented in Fig. 1.

In order to construct a shiftable multiscale and multidirectional decomposition, a combination of a multiresolution FB with the dual-tree DFB at high frequency has been proposed by Nguyen and Oraintara (2008). The combination of the multiresolution FB and the dual-tree DFB has a double objective. First, the multiscale FB provides a multiresolution image decomposition. Second, it can also remove the remaining aliasing components in the dual-tree DFB. In order to keep the complexity low, the FB should be efficiently implemented using separable filtering.

The proposed multiscale FB consists of an undecimated two-channel FB, and an iterated 2D multiresolution FB having decimation at the low-pass branch. Consider the construction in Fig. 1. At the front end, an undecimated two-channel FB ($L_0(z)$ and $R_0(z)$) is used to filter out the high-frequency components near $(\pm\pi, 0)$ and $(0, \pm\pi)$, which can potentially cause aliasing in the dual-tree. The high-pass filter produces a ‘residual’ image similar to that in a steerable pyramid (Simoncelli et al., 1992). It is clear that, for this undecimated FB to be PR, the filters must satisfy

$$|R_0(\omega)|^2 + |L_0(\omega)|^2 = 1. \quad (1)$$

The output of the wideband low-pass filter $L_0(z)$ is then fed into the first stage of the multiresolution pyramidal FB where the signal is divided into two parts: the coarse approximation (point L in Fig. 1) and the high-frequency component (point H in Fig. 1). This high-frequency component is then further decomposed by a dual-tree of DFBs to produce the real and imaginary parts of the 2^n complex directional subbands. This decomposition (block P) is also repeated at the low-frequency output (point L), forming a pyramid. The filters in blocks P and Q are designed to satisfy the PR and non-aliasing conditions:

$$|R_1(\omega)|^2 + \frac{1}{4}|L_1(\omega)|^2 = 1 \quad \text{and} \quad L_1(\omega_1, \omega_2) = 0, \quad (2)$$

where $\min(|\omega_1|, |\omega_2|) > (\pi/2)$, and the factor $1/4$ in the first equation is to ensure that the two-channel FB is PR.

The PDTDFB takes the shiftable pyramidal FB discussed above and applies a dual-tree 2^n -channel DFB at the high-pass output of each level. The first level of the PDTDFB is illustrated in Fig. 1. The blocks P and Q are iterated to provide a multiscale decomposition and synthesis. Based on the construction of the PDTDFB, the image representation produced by the structure in Fig. 1 is shiftable. The DFBs employed in the PDTDFB are constructed using the structure proposed in Do and Vetterli (2005). The primal and dual DFBs of the dual tree are identical at every level except for the second, where the filters of the two-channel fan FBs in the dual DFB have to satisfy the phase conditions. By this construction, the resulting directional filters in the primal and dual DFBs form Hilbert transform pairs. The conventional DFB is constructed by using a binary tree of two-channel FBs (Bamberger and Smith, 1992). These two-channel FBs can be obtained from one prototype fan FB if appropriate resampling blocks, which are upsampling or downsampling blocks having determinant one, are used in the tree (Park et al., 2004). The 2^n -channel dual-tree DFB is constructed from four-channel dual-tree DFBs by cascading two-channel FBs, using the same prototype fan FBs and resampling blocks. Therefore, each pair of directional filters in the primal and dual 2^n -channel DFBs still forms a Hilbert transform pair.

The image decomposition by the PDTDFB is thus a multiresolution and directional transform whose basis functions are the subband filters’ impulse responses. Therefore, the PDTDFB can be regarded as a shiftable contourlet transform (SCT). The SCT is shown to have a number of desirable properties for image analysis, including shiftable subband, arbitrarily high directionality, and low redundancy. Since the decimation ratio after $L_1(s)$ in each level

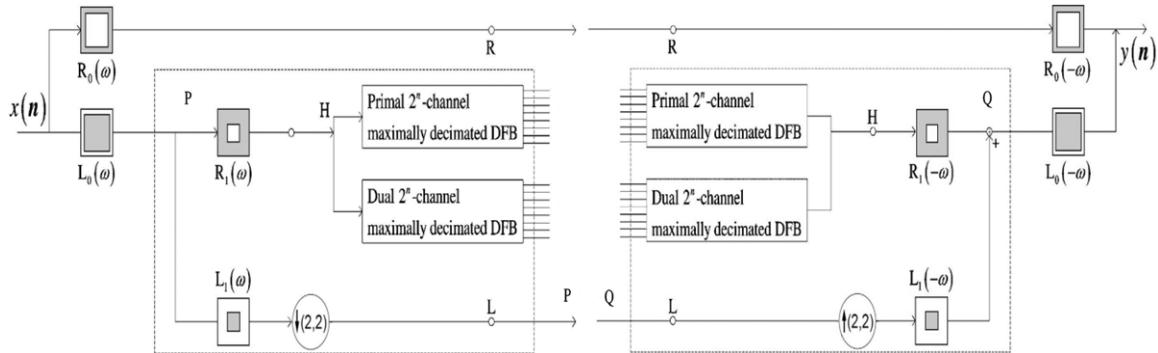


Fig. 1. The structure of the shiftable complex directional pyramid: analysis (left) and synthesis (right) sides (Nguyen and Oraintara, 2008).

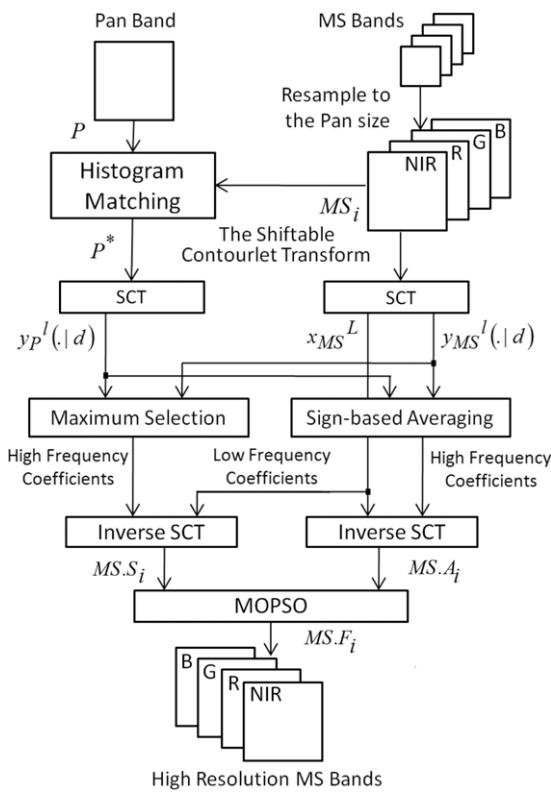


Fig. 2. Flowchart of the proposed pan-sharpening algorithm.

is $(1/4)$, the upper bound of the overcomplete ratio for the case of a single (maximally decimated) DFB is $1 + (1/4) + (1/6) = (4/3)$. Since there are two DFBs in each level, the overcomplete ratio is then bounded by $2 \times (4/3) = (8/3)$. Taking into account the residual subband at the output of filter $R_0(z)$, the total overcomplete ratio of the decomposition is $1 + (8/3) = (11/3)$ (Nguyen and Oraintara, 2008).

3. The proposed pan-sharpening algorithm

In this section, we describe the proposed pan-sharpening algorithm. Fig. 2 shows block diagram of the proposed method, which consists of a number of essential stages.

- (1) The input images are co-registered, and the MS data is resampled into the same spatial reference and grid as the pan band.
- (2) The pan and the MS images are histogram matched. This is because the final pan-sharpened image is less sensitive to the different illumination conditions in the pan image.

(3) In this step, an algorithm based on the shiftable contourlet transform is used to generate two initial results of the high-resolution MS images. The objective is to produce two modified high-resolution MS images, in which one has high spatial similarity to the pan image and the other one has high radiometric quality. Therefore, we have used two different fusion rules to integrate the high-frequency contourlet coefficients of the pan and MS images to generate two initial results of the high-resolution MS image or the pan-sharpened image.

(4) Finally, we can find an optimal pan-sharpened image by applying multiobjective particle swarm optimization and using the two initial pan-sharpened results. Specifically, the pan-sharpened image is obtained via a weighted combination of the two initial results, in which the weights are locally estimated via a multiobjective particle swarm optimization algorithm to generate a pan-sharpened image with high spatial and spectral qualities.

In the following subsections, we have provided a more detailed explication of the image fusion process.

3.1. Generation of two initial pan-sharpened results or the high-resolution MS images

Injection of the spatial details of the pan image into the MS image, without reducing the spectral quality or the color distortion of MS images were decreased, is the main challenge in the pan-sharpening methods. An ideal fusion method must be able to simultaneously deal with both spatial and spectral qualities, and not one at the cost of the other. In other words, the ideal fusion method should be able to preserve the original spectral and spatial information of the MS images while increasing the spatial resolution (Thomas et al., 2008). To reach this goal, we propose a two-step algorithm, in which two initial results of the high-resolution MS images are generated in the first step. Then, in the second step, the final pan-sharpened image is obtained through a weighted combination of the two initial results using an optimization algorithm.

The pan and MS images are locally different because of different resolutions, and therefore we cannot directly combine them in the spatial domain. For this reason, we generate two initial results, which are more appropriate for the weighted combination. We have used a multiresolution-based algorithm to generate the two initial results. In addition, the shiftable contourlet transform (SCT) is used for the multiresolution analysis. We show experimentally that the SCT has better transformation compared to the wavelet and contourlet transforms at the end of this section.

The first initial result is obtained using a simple maximum selection fusion rule in the shiftable contourlet transform domain. The objective is to generate a high-resolution MS image with high spatial similarity to the pan image. First, the pan and MS images are decomposed using the SCT. Since the high-frequency contourlet

coefficients reflect the image edges and detailed information, we have used the following formula to inject the detailed information or high-frequency coefficients of the pan image into the MS image:

$$y_1^l(.|d) = \begin{cases} y_p^l(.|d) & |y_p^l(.|d)| > |y_{MS}^l(.|d)| \\ y_{MS}^l(.|d) & \text{otherwise,} \end{cases} \quad (3)$$

where $y_1^l(.|d)$, $y_p^l(.|d)$, and $y_{MS}^l(.|d)$ are the high-frequency shiftable contourlet coefficients of the first high-resolution MS, the pan, and the low-resolution MS images, respectively. l is the decomposition level, d is the direction of contourlet coefficients in the transform domain, and $(.)$ is the shorthand notation of the spatial position.

In addition, the low-frequency shiftable contourlet coefficients of the first high-resolution MS image are selected from the low-resolution MS image. This is because the low-frequency coefficients contain the background information of an image, and therefore choosing background information from the low-resolution MS image in the first high-resolution MS image could preserve the color information. The inverse SCT of the combined contourlet coefficients generates the first high-resolution MS image (denoted by PS1):

$$PS1 = \Psi^{-1}(y_1, x_{MS}^L), \quad (4)$$

where Ψ is the inverse SCT, y_1 the high-frequency contourlet coefficients using (3), and x_{MS}^L the low-frequency contourlet coefficients of the low-resolution MS image in the last decomposition level L .

The second initial result is obtained using a new sign-based averaging fusion rule in the shiftable contourlet transform domain. The objective is to generate a high-resolution MS image with high radiometric quality, while the detailed information from the pan image is also available to some extent. The high-frequency shiftable contourlet coefficients of the second high-resolution MS image are obtained using the following formula:

$$y_2^l(.|d) = \begin{cases} \frac{y_p^l(.|d) + y_{MS}^l(.|d)}{2} & \text{si}(y_p^l(.|d)) = \text{si}(y_{MS}^l(.|d)) \\ \text{si}(y_p^l(.|d)) \times \frac{|y_p^l(.|d) - y_{MS}^l(.|d)|}{2} & \text{otherwise,} \end{cases} \quad (5)$$

where $y_2^l(.|d)$, $y_p^l(.|d)$, and $y_{MS}^l(.|d)$ are the high-frequency shiftable contourlet coefficients of the second high-resolution MS, the pan, and the low-resolution MS images, respectively. l is the decomposition level, d is the direction of contourlet coefficients in the transform domain, $(.)$ is the shorthand notation of the spatial position, and $\text{si}(.)$ is the sign function.

The simple averaging method produces a lot of near-zero coefficients, because the coefficients may be in different signs. The near-zero coefficients produce artifacts in the fused image after inverse transformation. In order to avoid this problem, we propose a new method for averaging the high-frequency coefficients in (5), which takes the sign of the coefficients into account. Similar to the first method, the low-frequency shiftable contourlet coefficients of the low-resolution MS image are used as low-frequency coefficients of the second high-resolution MS image. The inverse SCT of the combined contourlet coefficients generates the second high-resolution MS image (denoted by PS2):

$$PS2 = \Psi^{-1}(y_2, x_{MS}^L), \quad (6)$$

where Ψ is the inverse SCT, y_2 the high-frequency contourlet coefficients using (5), and x_{MS}^L the low-frequency contourlet coefficients of the low-resolution MS image in the last decomposition level L .

The two initial high-frequency MS results, which are obtained for four samples of pan and MS images, are shown in Fig. 3. As can be seen in Fig. 3, the first high-frequency MS image has high spatial quality because of transferring detailed information or high-frequency contourlet coefficients from both pan and

low-resolution MS images. In addition, the second high-resolution MS image, while it has detailed information from the pan image, is very close to the low-resolution MS image and has high radiometric quality. Spatial and radiometric qualities can be also measured by some objective functions. Indeed, there are many different ways to analyze the spatial and radiometric qualities. Here we have used the relative dimensionless global error in synthesis (ERGAS) and the relative average spectral error (RASE) as the radiometric and spectral qualities. The ERGAS calculates the amount of spectral and radiometric distortions in the image (Wald, 2000). The formula for the ERGAS is

$$\text{ERGAS} = 100 \frac{h}{l} \sqrt{\frac{1}{N} \sum_{n=1}^N \left(\frac{\text{RMSE}(n)}{\mu(n)} \right)^2}, \quad (7)$$

where $\frac{h}{l}$ is the ratio between pixel sizes of the pan and MS images, $\mu(n)$ is the mean of the n th band, and N is the number of bands. In addition, the root mean squared error (RMSE) is

$$\text{RMSE} = \sqrt{\frac{1}{L \times K} \sum_{i=1}^L \sum_{j=1}^K (F(i, j) - MS(i, j))^2}, \quad (8)$$

where F is the high-resolution fused image, MS is the original multispectral image, and $L \times K$ is the image size.

The RASE characterizes the average performance in the spectral bands (Ranchin and Wald, 2000):

$$\text{RASE} = \frac{100}{M} \sqrt{\frac{1}{N} \sum_{n=1}^N \text{RMSE}^2(n)}, \quad (9)$$

where M is the mean radiance of the N spectral bands of the original MS image.

The lower the values of the RASE and ERGAS indexes, the higher the radiometric and spectral qualities. To judge the spatial quality of the pan-sharpened image, we have used the method proposed by Zhou et al. (1998). First the pan and the high-frequency MS images are filtered using a Laplacian filter:

$$\text{Laplacian filter} = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}.$$

Then, we compare the resulting filtered images by considering the correlation coefficients between each band and the pan image. The closer the average correlation coefficient (CC) is to 1 indicates better spatial quality. We have obtained the three objectives (ERGAS, RASE, and CC to the pan) for the two initial pan-sharpened results of the high-resolution MS images shown in Fig. 3. As can be seen in Table 1, it is very clear that the first high-resolution MS image has high spatial similarity to the pan image and the second one has high radiometric quality. Indeed, the results show that the two initial results have the information we need in the next step. Specifically, in the next step, useful information within the two initial results is transferred into the final pan-sharpened image using a weighted combination scheme, in which the weights are optimally estimated using multiobjective particle swarm optimization.

3.2. Multiobjective optimization

Having the two initial results of the high-resolution MS images (PS1 and PS2), the final pan-sharpened (FPS) image is obtained using a weighted combination of the PS1 and PS2 images, in which the weights are optimally estimated via a multiobjective particle swarm optimization (MOPSO) algorithm to generate a pan-sharpened image with high spatial and radiometric qualities:

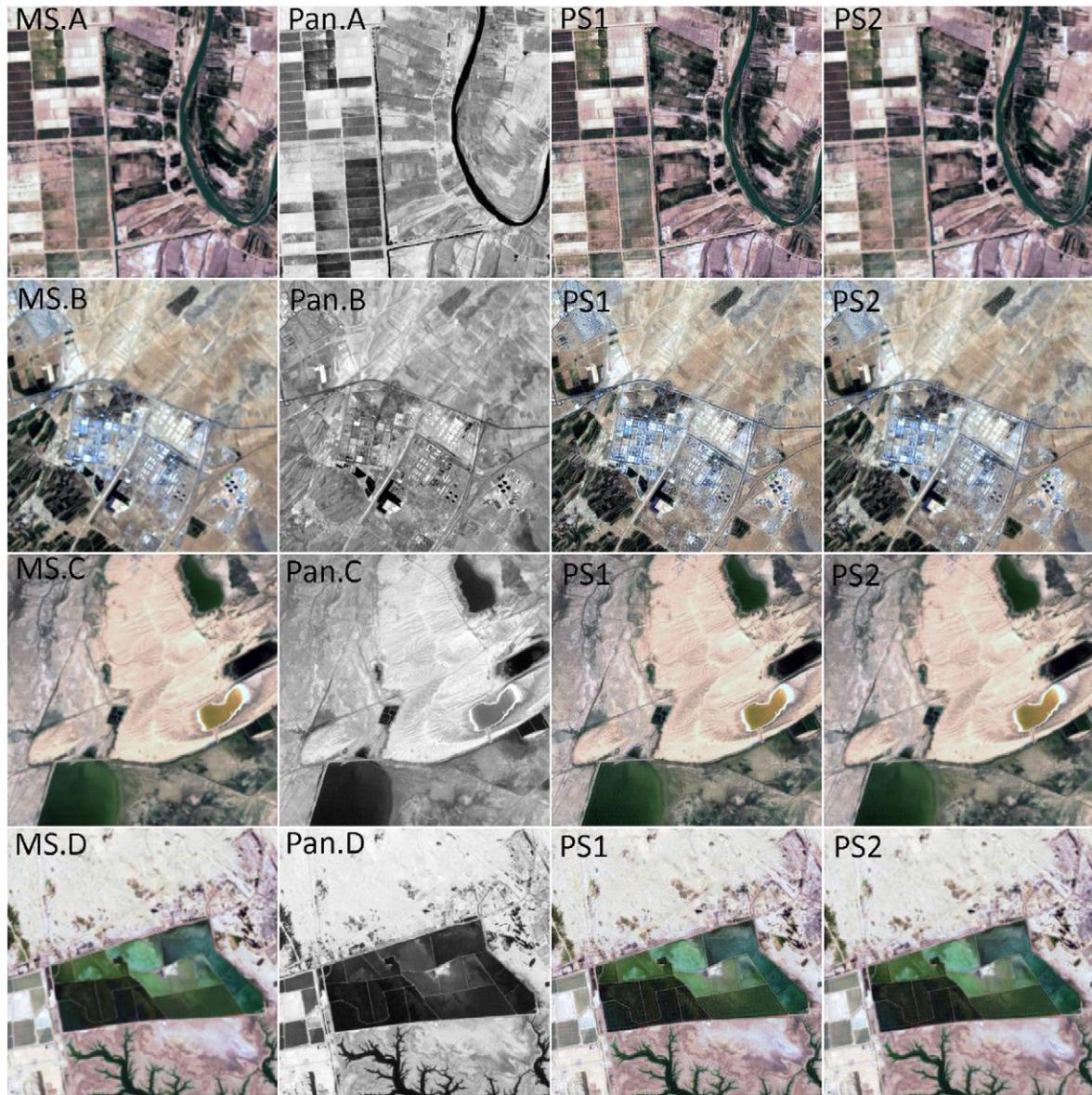


Fig. 3. True-color composites of the two initial pan-sharpened images for different multispectral and panchromatic images from the Landsat-7 dataset.

Table 1

Different objective measures for the two initial pan-sharpened results from the different datasets shown in Fig. 3.

Image	Index		
	Radiometric quality		Spatial quality
	ERGAS	RASE	CC to the pan
A	PS1	2.78	11.15
	PS2	1.14	4.53
B	PS1	2.77	11.14
	PS2	1.24	4.99
C	PS1	1.43	5.75
	PS2	0.97	3.74
D	PS1	1.52	6.08
	PS2	0.704	2.81

$$FPS_i = w_1 \times PS1_i + W2_i \times PS2_i \quad (10)$$

where FPS , $PS1$ and $PS2$ are the final pan-sharpened image and the two initial results, respectively. i is the number of bands of the MS image, $W1$ and $W2$ are the values in the range $[0,1]$, and $W1 + W2 = 1$. In addition, the weights are locally estimated using the MOPSO algorithm. The parameters $W1$ and $W2$ are not the same

dimensions as the $PS1$ and the $PS2$ images, and the \times operation does not represent Hadamard products between pairs of images. It should be mentioned that the parameters ($W1$ and $W2$) are separately obtained in each sliding window ($M \times N$) and each band of the two initial results. For an MS image with size $L \times K \times 4$, the number of parameters is $L \times K \times 4/M \times N$.

As we have mentioned, the final pan-sharpened image should have high spatial and spectral qualities. In addition, the spatial and radiometric qualities are measured by different indexes. Therefore, we have chosen two objectives for the optimization problem. In fact, the optimization problem is a biobjective optimization, which is a subset of multiobjective optimization.

The first objective is the correlation coefficient (CC) to the pan image, which measures the spatial quality of the pan-sharpened image. We have measured the CC to the pan in each band and each sliding window ($M \times N$) of the pan-sharpened (PS) image in (10) compared to the original panchromatic (Pan) image using the method proposed by Zhou et al. (1998). First, the pan and the PS images are filtered using the Laplacian filter. Then, we compare the resulting filtered images by considering the correlation coefficients

between each band of the *PS* and the pan images.

$$F_1 = CC(FPS, pan) = \frac{\sum_{m=1}^M \sum_{n=1}^N (PS_{n,m} - \bar{PS})(pan_{n,m} - \bar{pan})}{\sum_{m=1}^M \sum_{n=1}^N (PS_{n,m} - \bar{PS})^2 (pan_{n,m} - \bar{pan})^2} \quad (11)$$

where *FPS* and *pan* are the final pan-sharpened and original panchromatic images, and $M \times N$ is the size of the sliding window.

The second objective is the peak signal to noise ratio (PSNR), which measures the radiometric quality of the pan-sharpened image. As we have mentioned in the previous subsection, the ERGAS and RASE are the standard indexes to measure the radiometric and spectral qualities, but we cannot directly use them here. This is because the ERGAS and RASE are global indexes which are obtained using N bands of the pan-sharpened images compared to the original MS images, but here we measure the radiometric quality in each band of the pan-sharpened image compared to the corresponding band of the low-frequency MS image. Therefore, we have used the PSNR as the radiometric quality. We have measured the PSNR in each sliding window ($M \times N$) of the pan-sharpened (*PS*) image in (10) compared to the low-resolution MS image using the following:

$$F_2 = PSNR(FPS, MS) = 10 \log \frac{255^2}{RMSE^2} \quad (12)$$

in which the RMSE (root mean square error) is defined as

$$RMSE^2 = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N [FPS(n, m) - MS(n, m)]^2, \quad (13)$$

where *FPS* and *MS* are the final pan-sharpened and low-resolution MS images, and $M \times N$ is the size of the sliding window.

In other words, the optimization is a maximization problem, and the goal is to estimate the optimal parameters (*W*1 and *W*2) in each sliding window of the two initial results, which simultaneously maximize the two objectives expressed in (11) and (12). For this reason, we have used multiobjective particle swarm optimization, which will be explained in the next subsection.

3.3. Multiobjective particle swarm optimization (MOPSO)

Kennedy and Eberhart (1997) were the first to propose the particle swarm optimization (PSO) algorithm for optimization. The PSO algorithm is a population-based search algorithm based on the simulation of the social behavior of birds within a flock. In PSO, particles are flown through hyperdimensional search space. Changes to the position of the particles within the search space are based on the social-psychological tendency of individuals to emulate the success of other individuals (Sierra and Coello, 2006).

The position of each particle is changed according to its own experience and that of its neighbors. Let $x_i(t)$ denote the position of particle p_i at iteration t . The position of p_i is then changed by adding a velocity $v_i(t)$ to the current position; i.e.,

$$x_i(t) = x_i(t-1) + v_i(t). \quad (14)$$

The velocity vector reflects the socially exchanged information, and generally is defined in the following way:

$$v_i = W \times v_i(t-1) + C_1 r_1 \times (x_{pbest_i}(t) - x_i(t)) + C_2 r_2 \times (x_{leader}(t) - x_i(t)), \quad (15)$$

where *W* is the inertia weight employed to control the impact of the previous history of velocities on the current velocity of a given particle, and x_{pbest} is the personal best position of a given particle so far, that is, the position of the particle that has provided the

greatest success. x_{leader} is the position of the best particle of the entire swarm. C_1 and C_2 are learning factors, and they represent the attraction that a particle has toward either its own success or that of its neighbors, and $r_1, r_2 \in [0, 1]$ are random values.

In summary, the way in which the general single objective PSO algorithm works is as follows. First, the swarm is initialized. This initialization includes both positions and velocities. The corresponding *pbest* of each particle is initialized and the *leader* is located. Then, for a maximum number of iterations, each particle flies through the search space, updating its position using (14) and (15), and its *pbest* and, finally, the *leader* are updated, too.

In order to apply the PSO strategy for solving multiobjective optimization problems, it is obvious that the original scheme has to be modified. The solution set of a problem with multiple objectives does not consist of a single solution. Instead, in multiobjective optimization, we aim to find a set of different solutions (the so-called Pareto optimal set). There are many approaches to implement multiobjective problems using PSO. Here, we use Pareto approach, which consists of finding as many non-dominant solutions as possible and returning a set of non-dominant solution to the user (Sierra and Coello, 2006).

In this method, x_{pbest} and x_{leader} are updated at each iteration to minimize the distance of the Pareto front produced by the proposed algorithm with respect to the true (global) Pareto front (assuming that we know its location). In the proposed algorithm, x_{pbest} is updated at each iteration using the following:

$$x_{pbest_i}(t) = \begin{cases} x_i(t) & \text{if } F_1(x_i(t)) > F_1(x_i(t-1)) \\ x_{pbest_i}(t-1) & \text{and } F_2(x_i(t)) > F_2(x_i(t-1)) \\ & \text{otherswise,} \end{cases} \quad (16)$$

where x_{pbest} is the personal best position of a given particle, i is the particle number, t is the iteration, and F_1 and F_2 are the two objectives defined by (11) and (12).

In addition, a single *leader* is selected from the Pareto front solutions, which are stored in the external archive. In order to obtain Pareto front solutions, first the dominant solutions are selected between all the particles using the following:

$$x_d(i) = \begin{cases} x_i(t) & \text{if } F_1(x_i(t)) > F_1(x_i(t-1)) \\ & \text{and } F_2(x_i(t)) > F_2(x_i(t-1)) \\ \phi & \text{othersweise} \end{cases} \quad (17)$$

where $x_i(t)$ denotes the position of particle, i is the particle number, t is the iteration, and F_1 and F_2 are the two objectives defined by (11) and (12).

Having the dominant solutions, the Pareto front solutions are defined as follows:

$$PF(l)_{k=1,2,\dots,N,l \neq k} = \begin{cases} x_d(l) & \text{if } F_1(x_d(l)) > F_1(x_d(k)) \\ & \text{and } F_2(x_d(l)) > F_2(x_d(k)) \\ \phi & \text{othersweise,} \end{cases} \quad (18)$$

where $x_d(t)$ are the dominant solutions between all particles, and F_1 and F_2 are the two objectives defined by (11) and (12). $l = 1, 2, \dots, N$, and N is the number of dominant solutions.

As an example, Fig. 4(a) shows the particles at the first and second iterations, and Fig. 4(b) shows the dominant and Pareto front solutions. After finding the Pareto set, they are stored to an external archive, and then the external archive is updated with respect to the previous stored Pareto front set using the following:

$$PF^*(l)_{k=1,2,\dots,N,l \neq k} = \begin{cases} PF(l) & \text{if } F_1(PF(l)) > F_1(PF(k)) \\ & \text{and } F_2(PF(l)) > F_2(PF(k)) \\ \phi & \text{othersweise,} \end{cases} \quad (19)$$

where PF represents the Pareto front solutions defined by (18), and F_1 and F_2 are the two objectives defined by (11) and (12). $l = 1, 2, \dots, N$, and N is the number of Pareto front solutions in the external archive.

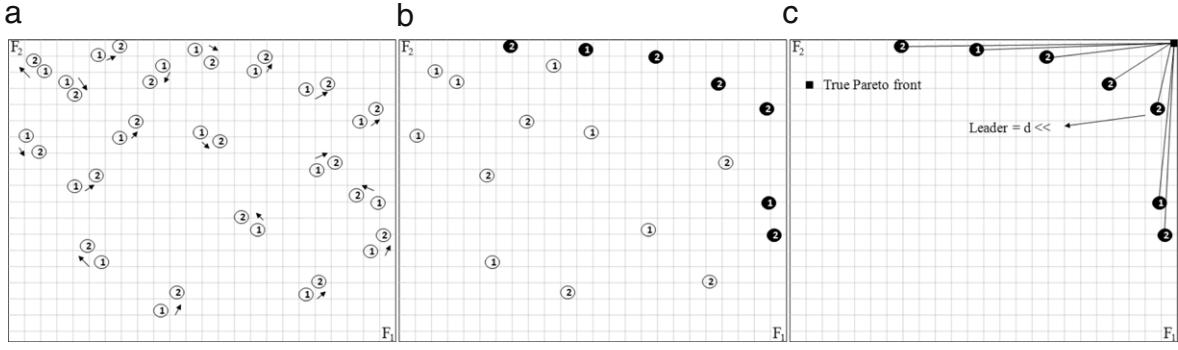


Fig. 4. (a) The particles at the first and second iterations, (b) the dominant (white circles) and Pareto front solutions (black circles), and (c) the true Pareto front, the Euclidian distances between the Pareto front set and the true Pareto front, and the leader.

Now, a single leader is chosen between the updated Pareto front set, which has minimum Euclidian distance to the true Pareto front. We have chosen the best values of the two objectives as the true Pareto front; they are obtained for the final pan-sharpened image produced by the proposed algorithm using (10) in each iteration of multiobjective optimization. As we have explained with respect to the optimization procedure, a number of particles are selected to estimate each parameter. Indeed, the true Pareto front is related to a particle, which results in the best values of the two objectives for the final pan-sharpened image:

$$TP_1 = \max_{j=1,2,\dots,NP} CC(FPS_j, pan) \quad (20)$$

$$TP_2 = \max_{j=1,2,\dots,NP} PSNR(FPS_j, MS), \quad (21)$$

where FPS is the final pan-sharpened image using (10), pan and MS are the original panchromatic and multispectral images, CC is the correlation coefficient defined by (11), and PSNR is the peak signal to noise ratio defined by (12). j is the particle number, NP is the number of particles, and TP_1 and TP_2 are the true Pareto front.

It should be mentioned that the true Pareto front are updated in each iteration. After determining the true Pareto front by (20) and (21), the *leader* is selected between the updated Pareto front set, which has minimum Euclidian distance to the true Pareto front:

$$d^2(i) = [F_1(PF^*(i)) - TP]^2 + \alpha^2 \times [F_2(PF^*(i)) - TP_2]^2. \quad (22)$$

Then,

$$\text{If } d(i) = \min_{j=1,2,\dots,K} d(j). \text{ Then Leader} = PF^*(i), \quad (23)$$

where PF^* are the updated Pareto front solutions, F_1 and F_2 are the two objectives defined by (11) and (12), TP_1 and TP_2 are the true Pareto front, K is the number of updated Pareto front solutions in the external archives, and $\alpha = TP_1/TP_2$. An example of the true Pareto front, the Euclidian distances between Pareto front set and the true Pareto front, and the *leader* is given in Fig. 4(c).

3.4. Parameter selection for the proposed pan-sharpening algorithm

In this subsection, the processing parameters for the proposed pan-sharpening algorithm are determined. The parameters consist of the decomposition level of the shiftable contourlet transform (SCT) to generate two initial pan-sharpened images, the block size of the sliding window for local parameter estimation in (10), and the PSO parameters selected for multiobjective optimization.

As we have mentioned, the model parameters (W_1 and W_2) in (10) are locally estimated over a sliding window. Therefore, the block size for local parameter estimation must be defined. Fig. 5(a) and (b) show the spatial and radiometric qualities of the final pan-sharpened image versus the block size for local

Table 2

The PSO parameters selected for multiobjective optimization.

Operation	Parameter
Particle number	15
Inertia weight	$W = 0.65$
Learning factors	$C_1 = 1, C_2 = 1$
Maximum iteration	40

parameter estimation. In addition, Fig. 5(c) shows the computation time of the proposed algorithm using different block sizes for multiobjective optimization. In this experiment, we have used the average results obtained for 20 MS (512×512) and corresponding pan (1024×1024) images from our Landsat-7 ETM + datasets. In addition, we have used two decomposition levels for the SCT, and the PSO parameters are manually set; see Table 2. It should be mentioned that the number of iterations and particles is set based on convergence of the optimization algorithm. More particles could speed up the convergence of the PSO algorithm; however, this increases the computation time. In addition, using too many particles and iterations is ineffective on the convergence of PSO. The inertia weight (w) is set to control the impact of the previous history of velocities on the current velocity of a given particle. High values of w produce a global search and low values of w produce a local search. Indeed, we want to consider both the local and global search properties of the inertia weights in the PSO algorithm, and therefore we give a medium value to it. In addition, the learning factors and the inertia weight are linearly decreased with respect to the iteration. In other words, we have used dynamic learning factors and inertia weight. This is because the larger inertia weights and the learning factors at the beginning help to find good seeds and the later small inertia weights facilitate a fine search.

As can be seen from Fig. 5(a) and (b), smaller block size results in a pan-sharpened image with higher spatial similarity (CC to the pan), and larger block size results in a pan-sharpened image with higher radiometric and spectral qualities (ERGAS, and RASE). In addition, the smaller block size increases the number of local parameters and the computation time. Therefore, there is a trade-off between the spatial and radiometric qualities of the final pan-sharpened image and the computation time of the proposed algorithm. Based on experiments, we have selected 35×35 block size for local parameter estimation, which results in acceptable spatial and radiometric qualities of the final pan-sharpened image and the computation time.

Another parameter is the number of decomposition levels of the shiftable contourlet transform (SCT) to generate the two initial pan-sharpened images in the proposed algorithm. Table 3 shows the decomposition levels used in the algorithm of the SCT versus the spatial and radiometric qualities of the final pan-sharpened image for different satellite imagery. In this experiment, we have

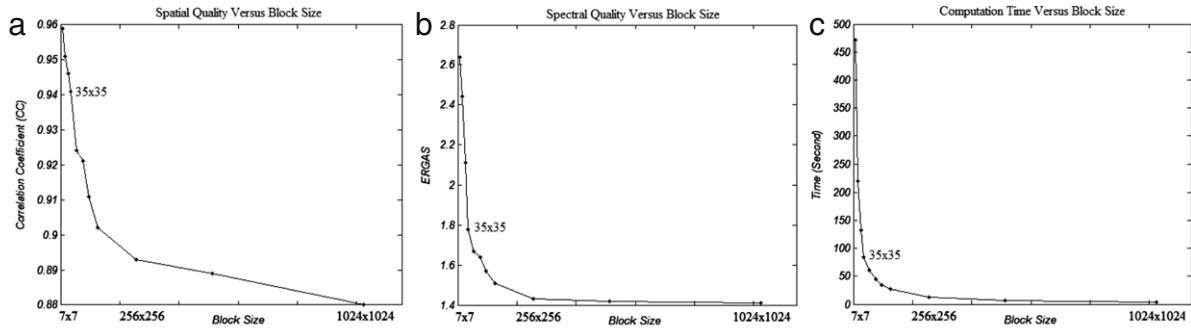


Fig. 5. (a) Overall CC to the pan index, (b) ERGAS index, and (c) computation time of the proposed algorithm versus the size of the sliding window for local parameter estimation.

Table 3

Average performance results obtained from different MS and pan images in the datasets using different decomposition levels of the SCT in the proposed algorithm.

Decomposition levels	CC to the pan			ERGAS			RASE		
	Landsat-7	Quickbird	Wordview2	Landsat-7	Quickbird	Wordview2	Landsat-7	Quickbird	Wordview2
1	0.901	0.897	0.889	1.26	1.19	1.05	5.09	5.20	4.83
2	0.947	0.928	0.921	1.78	1.57	1.23	7.28	6.87	6.29
3	0.951	0.951	0.949	2.76	2.34	2.85	11.09	9.11	9.72
4	0.957	0.957	0.951	3.24	3.97	4.13	13.04	13.79	14.11
5	0.962	0.961	0.958	3.64	4.32	4.56	14.64	16.14	16.56
6	0.968	0.969	0.963	4.12	4.98	5.04	15.53	17.11	17.34

used 35×35 block size for local parameter estimation. Table 3 shows that the decomposition level which maximizes both the spatial and the radiometric qualities of the final pan-sharpened image for the Landsat-7 dataset is two levels. Our explanation for this selection is that the gap for the spatial quality (CC to the pan) between the first and second decomposition levels is higher than the gap between the second and third decomposition levels, while the gap for the radiometric quality (ERGAS and RASE) between the second and third decomposition levels is much higher than the gap between the first and second decomposition levels. Using a similar judgment, three decomposition levels are selected for the Quickbird and Wordview2 satellite datasets.

3.5. Comparisons between different multiresolution transformations in the proposed algorithm

Another important aim of this paper is to show that the shiftable contourlet transform (SCT) is a better transformation approach than the contourlet transform (CT) and the dual-tree discrete wavelet transform (DT-DWT). As we mentioned above, the DWT suffers from shift variance and poor directionality, which are essential to the quality of wavelet-based image fusion (Rockinger, 1997). The DT-DWT is a modified version of the DWT, and it was proposed to overcome shift variance and directionality limitations of the DWT while maintaining a perfect reconstruction property with limited redundancy (Kingsbury, 2001). In addition, the contourlet transform (CT) is an alternative multiresolution method that provides an efficient directional representation and is efficient in capturing intrinsic geometrical structures of the natural image along the smooth contours (Do and Vetterli, 2005). Similar to the DWT, the CT is not shift invariant, and it results in artifacts along the edges to some extent. We mentioned in Section 2 that the SCT is a new shiftable and modified version of the CT. Indeed, the relation of the SCT to the CT is similar to that of the DT-DWT to the DWT. Therefore, the SCT is a better transformation than the CT because of the shift-invariant property, and it is better than the DT-DWT because of the directional selectivity.

First, we have objectively shown that the SCT performs better transformation compared to the CT and DT-DWT for transferring the spatial details of the pan image into the high-resolution MS

Table 4

Performance results obtained from 20 MS and pan images using different multiresolution transformations in the proposed algorithm.

Index	Multiresolution transformation		
	SCT	CT	DT-DWT
CC to the pan	Min	0.929	0.912
	Ave	0.947	0.933
	Max	0.958	0.948
	Std	0.0157	0.0190
ERGAS	Min	1.33	1.56
	Ave	1.78	2.43
	Max	2.01	2.91
	Std	0.389	0.754
RASE	Min	6.01	6.78
	Ave	7.28	8.45
	Max	8.53	10.21
	Std	1.26	1.71

image, without reducing the radiometric quality. Table 4 shows the objective results obtained from 20 MS and pan images in the database using different multiresolution transformations (DT-DWT, CT, and SCT) in the proposed algorithm.

In this experiment, we have used the results obtained for 20 MS (512×512) and corresponding pan (1024×1024) images from the Landsat-7 ETM+ datasets. We have used two decomposition levels for the DT-DWT, and two decomposition levels and 64 directions in each level of decomposition for the SCT and CT. For the DT-DWT-based method, the available "AntonB" mother wavelet is used. For the CT-based method, the available "9-7" is used for the pyramidal filter and "pkva" is used for the directional filter banks, and for the SCT-based method, "nalias" is used for the Laplacian low pass filter, "pkva" is used for the diamond or fan filter, and "9-7" is used for the wavelet filter.

For visual comparison, we have generated a pair of pan and MS images, in which the pan image is shifted up to two pixels in both the vertical and horizontal directions with respect to the MS image. Then we have used the DT-DWT and CT instead of the SCT in the proposed pan-sharpening algorithm to generate the two initial pan-sharpened images using (4) and (6). Fig. 6 shows the visual

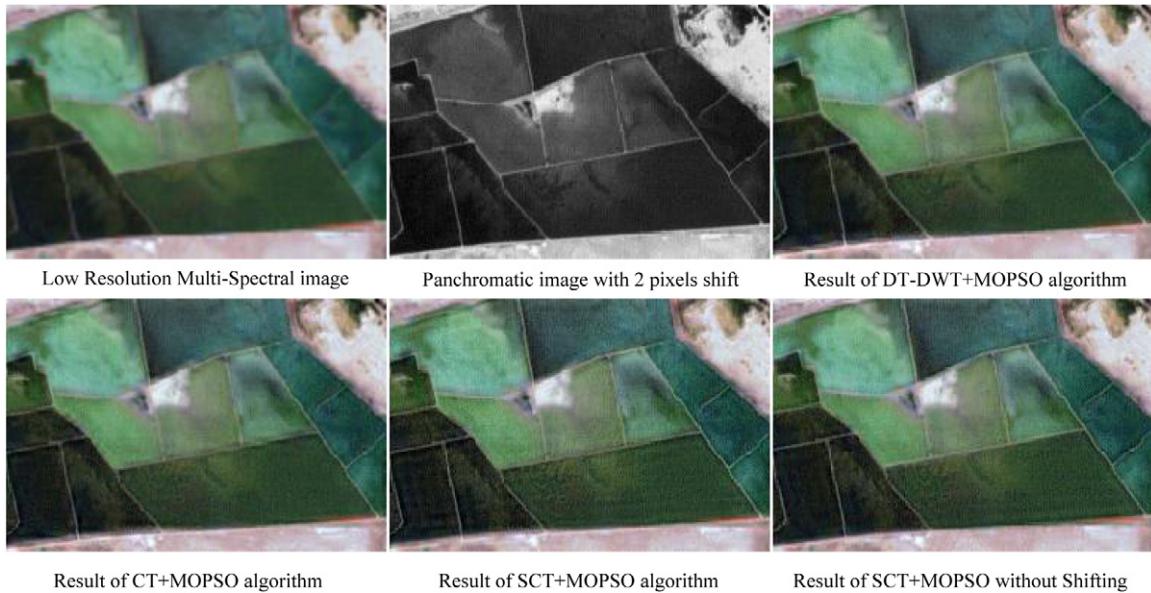


Fig. 6. Visual comparison between different multiresolution transformations in the proposed algorithm.

Table 5
Satellite dataset characteristics.

Features	Imagery type		
	Landsat-7 ETM +	Quickbird	Wordview-2
MS image-spectral bands and range	Blue: 450–515 nm Green: 525–605 nm Red: 630–690 nm Near-IR: 760–900 nm MID-IR 1: 1550–1750 nm MID-IR 2: 2080–2350 nm	Blue: 450–520 nm Green: 520–600 nm Red: 630–690 nm Near-IR: 760–900 nm	Coastal: 400–450 nm Blue: 450–510 nm Green: 510–580 nm Yellow: 585–625 nm Red: 630–690 nm Red Edge: 705–745 nm Near-IR1: 770–895 nm Near-IR2: 860–1040 nm
Pan image-spectral range	520–920 nm	450–900 nm	450–800 nm
Spatial resolution-MS image	28.5 m	2.4 m	1.84
Spatial resolution-pan image	14.25 m	0.60 m	0.46 m
Resolution ratio of MS and pan images	2	4	4

Table 6
Dataset summary.

Features	Imagery type		
	Landsat-7 ETM+	Quickbird	Wordview-2
Location	Tehran, Iran	Boulder, CO, USA	Kokilai Lagoon, Sri Lanka
MS image size	7348 × 6208	828 × 815	3198 × 2667
Pan image size	14 696 × 12 416	3312 × 3260	12 792 × 10 668

comparison of the different methods. It can be seen in Fig. 6 that using the SCT in the proposed algorithm results in a pan-sharpened image with higher spatial and radiometric qualities compared to the other transformations. Using the CT in the proposed algorithm causes artifacts around the edges. Our explanation for this is that the SCT is a shiftable transform, while the CT is not shift invariant, and therefore a shift-dependent fusion method leads to unstable and flickering results. The human visual system is primarily sensitive to moving light stimuli, so moving artifacts, when introduced by the fusion process, are highly distracting to a human observer (Rockinger, 1997). In addition, the detailed information of the SCT + MOPSO approach is better compared to the DT-DWT + MOPSO, which indicates that contourlets are better than wavelets for spatial detail injection. This is because the DT-DWT has only six directional subbands, while we have used 64 directions in the SCT method in each level of decomposition.

4. Experimental results and datasets

4.1. Experimental datasets

The presented method is evaluated by performing pan-sharpening on datasets acquired by Landsat-7 ETM+, Quickbird, and Wordview-2 satellites. All images are radiometrically calibrated and orthorectified to uniform ground resolution. The MS and pan images are co-registered for each dataset. All imageries have different characteristics in terms of spectral bands and range, and spatial resolutions, etc. These characteristics are summarized in Table 5. Table 6 presents a summary of the datasets as related to their size and location. The true-color composite sample imageries used for the experiments are shown in Fig. 7.

- (1) The Landsat-7 ETM + dataset is shown in Fig. 7(a), which covers different urban, forested, Mediterranean, and agricultural

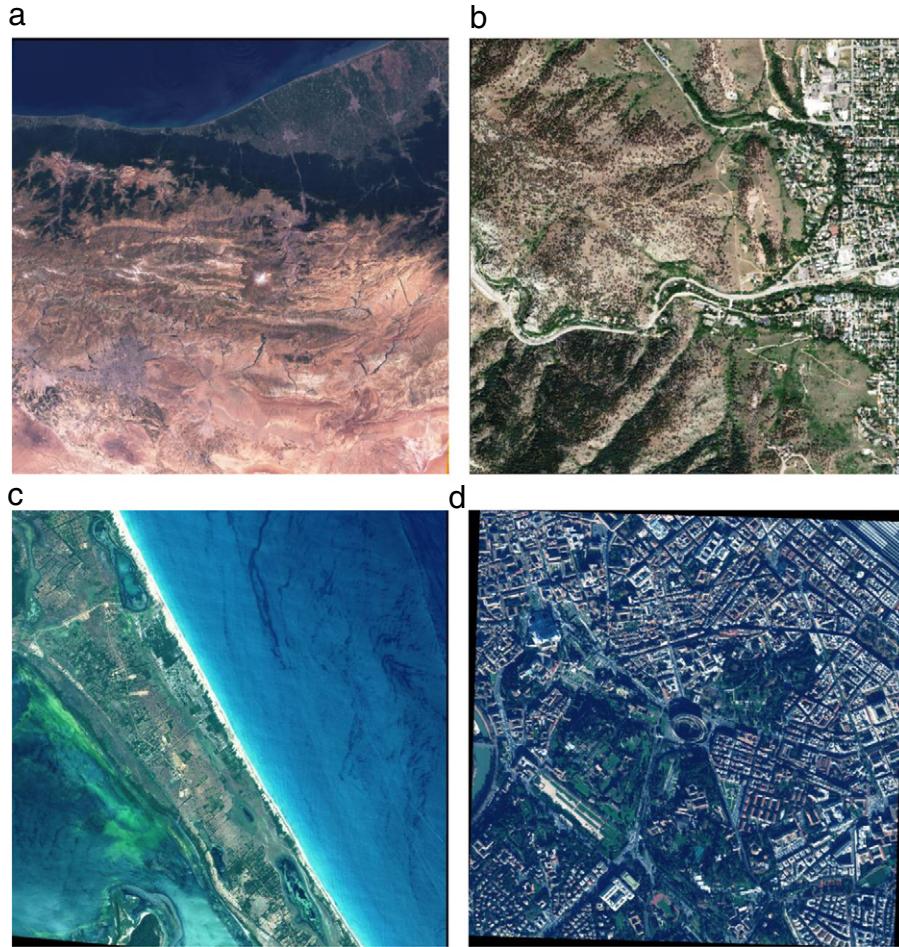


Fig. 7. True-color composites of multispectral data used in the experiment. (a) Lansat-7 ETM + scene from Iran, (b) Quickbird scene from Boulder, CO, USA, (c) Quickbird scene from Kokilai Lagoon, Sri Lanka, and (d), Wordview2 scene from Rome, Italy.

- areas. Because only three bands are within the wavelength span of the pan image, only the green, red, and near-infrared (near-IR) bands are considered for pan-sharpening.
- (2) The second dataset is very-high-resolution image data collected by the Quickbird satellite. Two QuickBird imageries are shown in Fig. 7(b) and (c). The first dataset, shown in Fig. 7(b), is an imagery representing mountainous regions consisting of forest and having a lot of uniform texture details, and urban features such as manmade buildings. The second dataset, shown in Fig. 7(c), covers different Mediterranean and agricultural areas. The four MS bands of the QuickBird span the visible and near-IR wavelengths are within the wavelength span of the pan image, and we have considered all of them for pan-sharpening.
 - (3) The third very-high-resolution dataset considered, which has been acquired by the Wordview2 satellite, is shown in Fig. 7(d). This imagery represents urban features such as manmade buildings. We have considered the four MS bands consisting of the blue, green, yellow, and red bands for pan-sharpening. It is possible to pan-sharpen the red-edge band, because it is also within the wavelength span of the pan image.

In addition, for simplicity of working with the images and pan-sharpening algorithms, we have generated 20 MS (512×512) and corresponding pan (1024×1024) images from the Lansat-7 ETM + dataset, 10 MS (256×256) and corresponding pan (1024×1024) images from the Quickbird datasets, and 10 MS (256×256) and corresponding pan (1024×1024) images from the Wordview2 dataset. The pan and MS images are co-registered

for each dataset, and the MS images were spatially resampled by a factor of 2 for the Lansat-7 ETM + dataset, and by a factor of 4 for the Quickbird and Wordview2 datasets, to produce new 14.25 m, 0.6 m, and 0.46 m resolution MS images, respectively.

4.2. Experimental results

There are many different ways to analyze the results of pan-sharpened images and to compare different methods. When comparing different methods, we are interested in spatial, radiometric, and spectral qualities. In judging spatial quality, it is relatively easy to check the sharpness of the edges. However, when judging spectral and radiometric qualities, it is much more difficult to match the colors of the result to the original multispectral ones by visual inspection. Many metrics analyze the spectral quality. Here, we have used four metrics to judge the spectral and radiometric qualities: the relative dimensionless global error in synthesis (ERGAS) proposed by Wald (2000), the relative average spectral error (RASE) (Ranchin and Wald, 2000), the spectral angle mapper (SAM) (Yuhas et al., 1992), and the Q_4 index (Alparone et al., 2004). The Q_4 quality index is a generalization of the universal image quality index (Q_{avg}) (Wang and Bovik, 2002). The RASE, ERGAS, and Q_4 indexes measure both the radiometric and spectral qualities, and the SAM metric only measures the spectral quality.

To judge the spatial quality of the pan-sharpened image, we have used the correlation coefficient (CC) between the pan and pan-sharpened images (Zhou et al., 1998). Here, we compare the spectral and radiometric qualities (ERGAS, RASE, SAM, and Q_4

Table 7

Average performance of different state-of-the-art pan-sharpening algorithms against the proposed method for different datasets.

Index	Dataset	Pan-sharpening algorithms						RV
		IHS	Brovey	PCA	P + XS	Ehlers	BDF	
CC to the pan	Landsat-7	0.937	0.967	0.961	0.880	0.931	0.963	0.947
	Quickbird	0.915	0.971	0.974	0.792	0.884	0.971	0.951
	Wordview2	0.923	0.967	0.916	0.743	0.898	0.963	0.949
ERGAS	Landsat-7	5.55	6.16	6.38	2.51	2.96	3.23	1.78
	Quickbird	10.98	9.86	9.71	2.88	3.85	4.84	2.34
	Wordview2	6.94	9.96	9.87	3.43	3.89	5.48	2.85
RASE	Landsat-7	22.10	27.20	25.56	10.30	11.87	13.23	7.28
	Quickbird	43.00	40.62	36.92	11.23	15.02	19.15	9.11
	Wordview2	23.17	32.18	29.54	11.14	12.92	19.38	9.72
SAM	Landsat-7	1.69	0.0	2.35	1.70	1.41	1.99	1.01
	Quickbird	3.55	0.0	4.63	2.36	2.21	3.15	2.14
	Wordview2	2.65	0.0	6.34	2.52	2.37	3.11	1.69
Q ₄	Landsat-7	0.928	0.891	0.848	0.904	0.927	0.881	0.925
	Quickbird	0.893	0.843	0.819	0.921	0.939	0.895	0.943
	Wordview2	0.941	0.899	0.893	0.949	0.964	0.928	0.971

indexes) of different pan-sharpened results with respect to their spatial quality (CC to the pan index). In other words, the algorithm which generates a better trade-off between the spatial, radiometric and spectral qualities is chosen as the best method, and higher spatial or radiometric or spectral quality is individually ineffective.

It should be mentioned that to assess the objective or numerical results (ERGAS, RASE, SAM, and Q₄ indexes) of the pan-sharpened images, spatially degraded pan and MS images from the original datasets are used, and the results obtained are compared to the original data. In addition, three subsets of the original (not degraded) datasets were extracted to compare the visual effect of different pan-sharpening algorithms. (It is more usual to use real data for visual comparison.) Therefore, there is no reference high-resolution MS image in the visual or subjective comparisons. The results will be discussed in the following two subsections. First, we demonstrate the results of standard pan-sharpening algorithms. Then, the results of multiresolution-based methods will be illustrated.

4.2.1. Comparisons between different state-of-the-art pan-sharpening algorithms

In this subsection, we compare the proposed method, which is based on shiftable contourlet transform and multiobjective particle swarm optimization, with some of the best state-of-the-art techniques consisting of the IHS, Brovey (Pohl and Genderen, 1998), PCA,¹ P + XS¹ (Ballester et al., 2006; Ehlers, 2004), and Bayesian Data Fusion (BDF)² (Fasbender et al., 2008). The parameters of each method have been set according to the values given by their respective authors in their corresponding papers.

Table 7 summarizes the results obtained. As can be seen in Table 7, the proposed method provides better spectral and radiometric qualities (ERGAS, RASE, SAM, and Q₄ indexes) in all cases and better spatial quality (CC to the pan index) in almost all cases. When looking closer at the results, we observe the following.

- The proposed method outperforms the IHS technique in spatial, radiometric and spectral qualities. The CC index of the IHS method is acceptable, while the corresponding spectral and radiometric qualities are very weak.
- Our method gives better results than the Brovey algorithm for radiometric quality. However, the Brovey method gives a better result for the CC to the pan index, but the corresponding ERGAS, RASE, and Q₄ values for it are very weak, which indicates

that the pan-sharpened image has low radiometric quality. It should be mentioned that the angle from the spectral angle mapper (SAM) between the original and pan-sharpened MS images is zero. This is because the operation of multiplication only changes the norm of a spectral vector (Du et al., 2007). Therefore, it does not affect the SAM index between the original and pan-sharpened MS images. A value of SAM equal to zero denotes absence of spectral distortion, but radiometric distortion is possible, and each band of the MS image is greatly changed due to different values being involved in the multiplication from pixel to pixel (the two pixel vectors are parallel but have different lengths).

- Our method provides better results than the PCA algorithm. While the CC index of our method has a little difference with that of the PCA method, for other indexes our method produces a higher gap with those of the PCA.
- The proposed method outperforms the P + XS and Ehlers techniques for spatial, radiometric and spectral qualities for different datasets.
- Our method gives better results than the BDF algorithm for spectral and radiometric quality. The BDF method provides very good spatial quality, while the corresponding spectral and radiometric quality indexes are weak.

For subjective evaluation of methods, there are two important criteria: spectral and spatial qualities. In other words, the final pan-sharpened image, at the same time as providing high spatial quality (or spatial details from the panchromatic image), should have high spectral and radiometric qualities (or color information from the original multispectral image).

Figs. 8 and 9 give subjective results for two samples of pan and MS images from the Wordview2 and Quickbird datasets, respectively. The two samples are selected from real (not degraded) datasets (Table 6), and therefore there are no reference images for comparison. (The resolution ratio between the MS and pan images is 4.) It can be seen in Figs. 8 and 9 that the results of the IHS and Brovey methods are sharp enough, but the colors are changed compared to the low-resolution MS images. The result of the PCA method is not clear, and the colors are changed. With the P + XS method, the colors are preserved, but the spatial details from the pan image are not appropriately transferred to the pan-sharpened image. The Ehlers method suffers from a misregistration problem. This is because, in the Ehlers method, a low-frequency version of the intensity image from the MS bands is added to the high-frequency version of the pan image, and if the pan and MS images are not accurately registered, this causes an unclear image with thick edges. In addition, the Ehlers method uses a Fourier transform for spatial detail extraction from the pan image, and the Fourier

¹ Available at: www.math.ucla.edu/~wittman/pansharpening/index.html.

² ORFEOT Accompaniment Program, ORFEOT Tool Box, 2007. Available: http://smc.cnes.fr/PLEIADES/A_prog_accomp.htm.

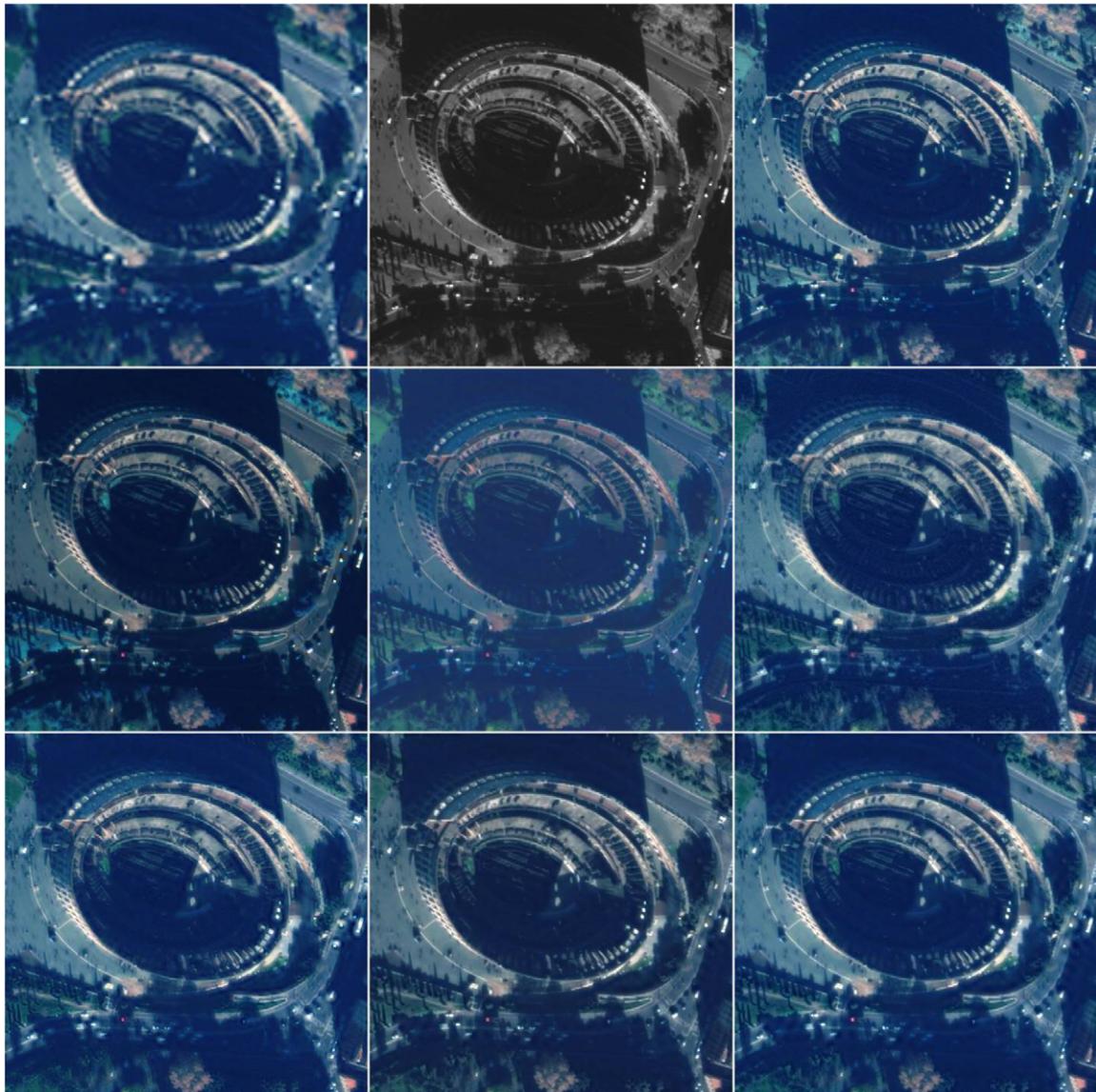


Fig. 8. Subjective results of different pan-sharpening methods. True-color composites of a low-resolution multispectral image from the Wordview2 dataset, the corresponding panchromatic image, and the result of the IHS method (top row, left to right). The results of the Brovey, PCA, and P + XS methods (middle row). The results of the Ehlers, BDF, and the proposed methods (bottom row).

transform is not an efficient representation for 2D signals. The BDF method results in very good spatial quality in the pan-sharpened image, and detail information from the pan image is appropriately preserved, while the spectral and radiometric information from the low-resolution MS image has changed to some extent. The result of the proposed algorithm visually demonstrates very good spatial quality and the colors are well preserved from the low-resolution MS image. Indeed, a visual inspection of the true-color composites confirms the numerical results of Table 7, in which the proposed method gives a better trade-off between the spatial and spectral qualities.

4.2.2. Comparisons between different multiresolution-based pan-sharpening algorithms

In this experiment, we have compared the proposed algorithm, which is based on the shiftable contourlet transform (SCT) and the multiobjective particle swarm optimization (MOPSO) algorithm, with multiresolution-based pan-sharpening methods. In the multiresolution-based methods, first the MS and pan images are decomposed. Then, the high-frequency coefficients of the

final pan-sharpened image are obtained using averaging (A) and substitutive (S) fusion rules. The final pan-sharpened image is obtained using an inverse multiresolution transform. Here we have used DT-DWT³ and CT⁴ transforms for the multiresolution analysis: DT-DWT + A , DT-DWT + S (Ioannidou and Karathanassi, 2007), CT + A , and CT + S . In addition, we have compared the proposed method with two hybrid multiresolution algorithms. In the first hybrid method, the IHS transform is used to separate the spectral information of the low-resolution MS image, and then the multiresolution-based algorithm is used to inject detailed information from the histogram-matched pan image to the intensity image. Finally the inverse IHS transform generates the pan-sharpened image: IHS + DT-DWT + S (Zhang and Hong, 2005), and IHS + CT + S (Hui and Cheng, 2008). In the second hybrid method, contourlet and principal component analysis (PCA)

³ Available at: taco.poly.edu/WaveletSoftware/dt2D.html.

⁴ Available at: www.mathworks.com/matlabcentral/fileexchange/8837-contourlet-toolbox.

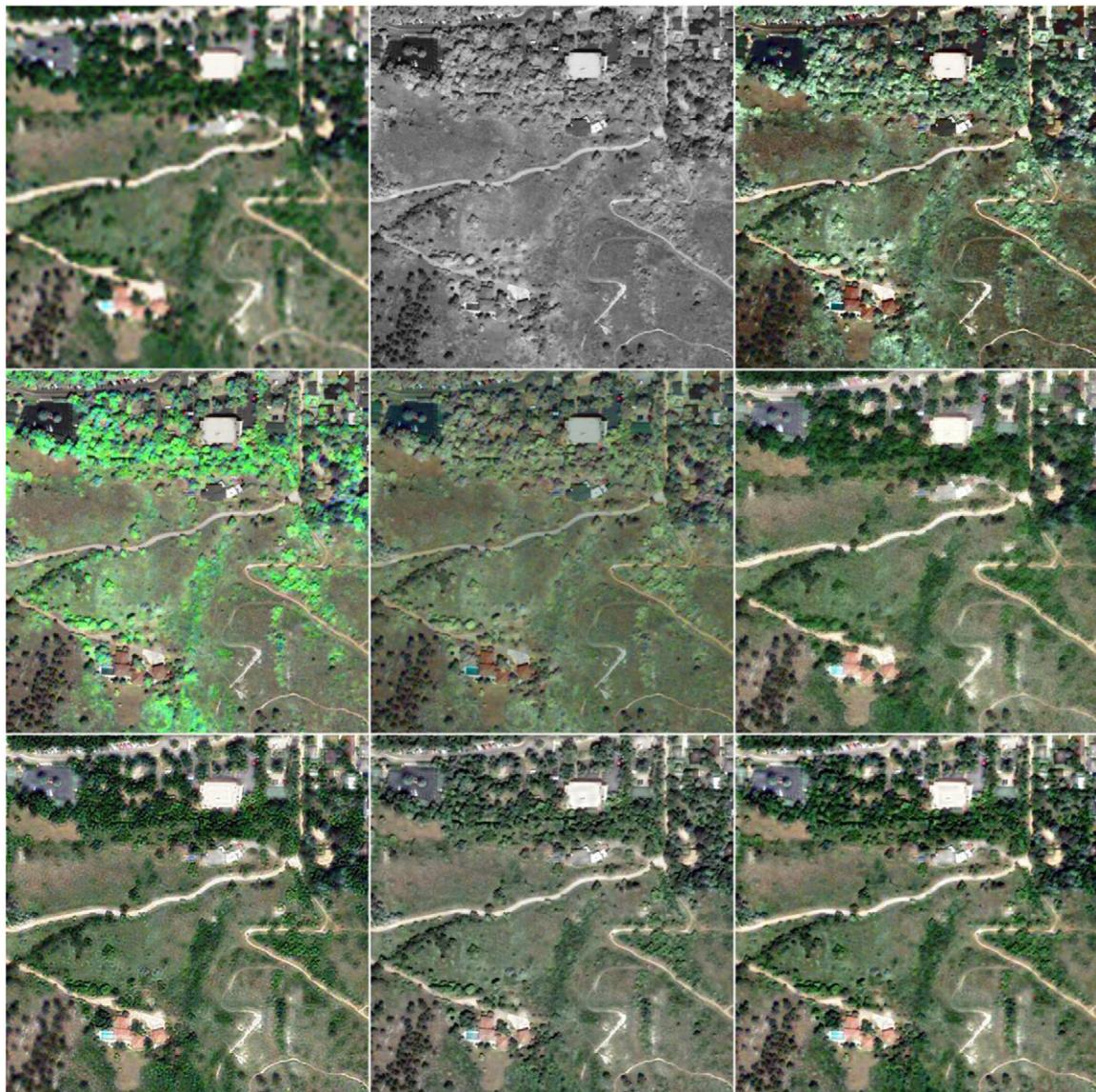


Fig. 9. Subjective results of different pan-sharpening methods. True-color composites of a low-resolution multispectral image from the Quickbird dataset, the corresponding panchromatic image, and the result of the IHS method (top row, left to right). The results of the Brovey, PCA, and P + XS methods (middle row). The results of the Ehlers, BDF, and the proposed methods, (bottom row).

methods have been used for spatial and spectral transformations, respectively. Then, the first principal component of high variance is used for replacing or injecting it with high spatial details from the high-resolution histogram-matched pan image. Injecting of the spatial details is based on the multiresolution-based algorithm, in which the high-frequency contourlet coefficients of the pan image are replaced with the high-frequency contourlet coefficients in the first principal component. Finally, the inverse contourlet and PCA transformations generate the final pan-sharpened image (Vijay et al., 2008).

Table 8 summarizes the results obtained for different datasets. As can be observed in Table 8, the proposed method provides appropriate spectral and radiometric qualities (ERGAS, RASE, SAM, and Q_4 indexes) with respect to the spatial quality (CC to the pan index) compared to other methods. When looking more closely at the results, we observe the following.

- The CC to the pan index for methods which use the substitutive fusion rule (S) are better than for other algorithms, while the spectral and radiometric quality indexes for them are worse.

This indicates that the pan-sharpened images using these methods are sufficiently sharp, but the spectral information from the low-resolution MS image is not appropriately transferred to the pan-sharpened image.

- The spectral and radiometric quality indexes for methods which use the averaging fusion rule (A) are better than for other algorithms, while the corresponding CC to the pan index for them is very low. This indicates that the colors from the low-resolution MS image are well preserved, but the image is not clear or sharp.
- We have discussed above that the shiftable contourlet transform is a better transformation than the CT because of its shift-invariant property, and is better than the DT-DWT because of its directional selectivity. The objective results demonstrate that the SCT is a better transformation compared to the DT-DWT and CT for both substitutive and averaging fusion rules.
- The results of using the IHS transform as a spectral transformation in the multiresolution-based scheme show better spectral and radiometric qualities compared to the multiresolution-based methods which use substitutive fusion rule (S).

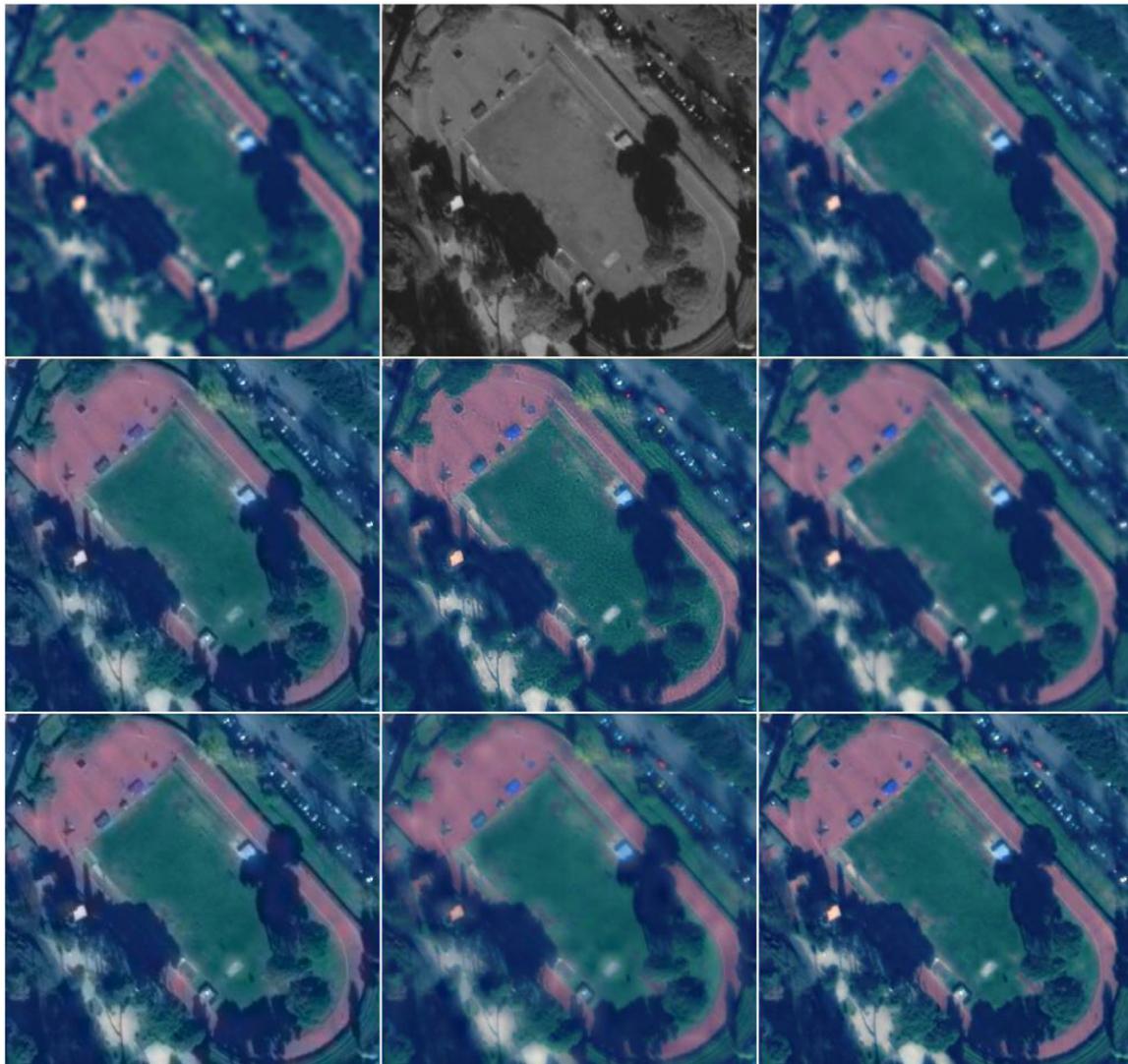


Fig. 10. Subjective results of different multiresolution-based pan-sharpening methods. True-color composites of a low-resolution multispectral image from the Wordview2 dataset, the corresponding panchromatic image, and the result of the DT-DWT + A method (top row, left to right). The results of the DT-DWT + S, IHS + DT-DWT + S, and CT + A methods (middle row). The results of the CT + S, PCA + CT, and proposed methods (bottom row).

Table 8

Average performance of different multiresolution-based pan-sharpening algorithms against the proposed method for different datasets.

Index	Dataset	Pan-sharpening algorithms										RV
		DTDWT + A	DTDWT + S	CT+A	CT+S	IHS+DTDWT+S	IHS+CT+S	PCA+CT	SCT+S	SCT+A	MOPSO	
CC to the pan	Landsat-7	0.906	0.971	0.912	0.976	0.961	0.964	0.921	0.983	0.915	0.947	1
	Quickbird	0.907	0.964	0.905	0.963	0.960	0.957	0.915	0.967	0.909	0.951	
ERGAS	Wordview2	0.889	0.967	0.884	0.964	0.963	0.961	0.912	0.972	0.891	0.949	0
	Landsat-7	1.34	3.68	1.21	3.37	3.27	3.21	1.85	3.08	1.16	1.78	
RASE	Quickbird	1.83	3.59	1.96	3.84	3.45	3.67	2.70	3.42	1.81	2.34	0
	Wordview2	2.19	4.26	2.27	4.41	4.14	4.23	3.04	4.18	2.13	2.85	
SAM	Landsat-7	5.40	13.76	5.21	13.21	12.32	12.03	7.43	12.11	4.97	7.28	0
	Quickbird	7.12	13.95	7.60	14.89	13.51	14.37	10.27	13.67	7.01	9.11	
Q ₄	Wordview2	7.23	14.21	7.51	14.63	13.62	14.09	10.13	14.17	7.18	9.72	1
	Landsat-7	0.728	1.43	0.794	1.57	1.47	1.46	1.37	1.41	0.715	1.01	
Quickbird	1.47	3.10	1.62	3.38	2.77	2.86	2.57	3.01	1.42	2.14	0	1
	Wordview2	1.23	2.27	1.26	2.42	2.18	2.21	1.79	2.13	1.11	1.69	
Wordview2	Landsat-7	0.928	0.838	0.933	0.826	0.924	0.921	0.928	0.859	0.941	0.925	1
	Quickbird	0.951	0.881	0.942	0.864	0.931	0.927	0.939	0.887	0.954	0.943	
Wordview2	Wordview2	0.969	0.956	0.971	0.942	0.952	0.944	0.965	0.961	0.983	0.971	1

- The PCA + CT method, which uses spatial and spectral transformations, results in a better trade-off between the spatial

and spectral qualities compared to the multiresolution-based methods which use the averaging fusion rule (A). However, the

- spatial quality for different datasets is very low, which is the consequence of using spectral transformation (PCA) in the pan-sharpening scheme.
- The proposed pan-sharpening method produces a better trade-off between the spatial (CC to the pan index), radiometric and spectral qualities (ERGAS, RASE, SAM, and Q_4 indexes) compared to other techniques. In other words, the results show that the proposed method could transfer spatial details of the pan image into the final pan-sharpened image without reducing the spectral and radiometric qualities of the MS image.

In addition, Fig. 10 gives subjective results for a sample of pan and MS images from the Wordview2 datasets. The sample are selected from real (not degraded) datasets (Table 6), and therefore there is no reference image for comparison. (The resolution ratio between the MS and pan image is 4.)

It can be seen in Fig. 10 that the colors in the results of the CT + A and DT-DWT + A methods are preserved, while the images are not sharp. The results of the CT + S, and DT-DWT + S methods are sufficiently sharp, while the colors (especially the blue, orange, and red areas from the low-resolution MS image) are changed. The result of the IHS + DT-DWT + S method locally introduces small artifacts due to adding inaccurate spatial details of the low-resolution MS image to the pan image. The IHS + CT + S method also results in a similar pan-sharpened image. The result of the PCA + CT method is not very accurate in spectral or color preservation, and the spatial information of the pan image is not appropriately transferred into the pan-sharpened image. The result of the proposed algorithm visually demonstrates very accurate spectral preservation and superior performance in terms of radiometric and geometric accuracy. Visual assessment of the true-color composites confirms the numerical results of Table 8, in which the proposed method provides a better trade-off between the spatial and spectral qualities.

5. Conclusions

This paper presents a novel pan-sharpening algorithm based on the shiftable contourlet transform and multiobjective particle swarm optimization. The shiftable contourlet transform is known to provide a shiftable transform compared to the contourlet transform, and better directional representation compared to the wavelet transform. In addition, multiobjective particle swarm optimization is used to generate high-resolution MS images with high spatial similarity to the pan image and high radiometric quality in each band. Therefore, as expected, the proposed method provides better pan-sharpened results based on the global metrics (CC to the pan, ERGAS, RASE indexes) which measure the spatial and radiometric qualities. The experimental results also show that the pan-sharpened images produced have good spectral quality based on the Q_4 and SAM metrics.

One aspect that we would like to explore in the future is to analyze the evaluation index system to acquire a meaningful measurement, and to use a better combination of the evaluation indexes in the multiobjective optimization procedure. We are also considering improving the optimization performance by using other evolutionary algorithms such as the genetic algorithm. Using a segmentation map of the source images instead of a rectangular window for local parameter estimation in the proposed method can improve the pan-sharpened results. Future work should be done on these promising issues.

Acknowledgements

The authors would like to express their great appreciation to the Iranian Space Agency for providing the Landsat-7 + ETM images for this research. In addition, the QuickBird, and Wordview2 imagery were obtained from Space Imaging, DigitalGlobe, and

Global Land Cover Facility (GLCF), respectively. The valuable comments of the anonymous reviewers obviously contributed to the improvement of this paper.

References

- Aiazz, B., Alparone, L., Baronti, S., Garzelli, A., 2002. Context-driven fusion of high spatial and spectral resolution images based on over sampled multiresolution analysis. *IEEE Transaction on Geoscience and Remote Sensing* 40 (10), 2300–2312.
- Alparone, L., Baronti, S., Garzelli, A., Nencini, F., 2004. A global quality measurement of pan-sharpened multispectral imagery. *IEEE Geoscience and Remote Sensing Letters* 1 (4), 313–317.
- Amolins, K., Zhang, Y., Dare, P., 2007. Wavelet based image fusion techniques—an introduction, review and comparison. *ISPRS Journal of Photogrammetry and Remote Sensing* 62 (4), 249–263.
- Audicana, M.G., Saleta, J.L., Catalan, O.G., Garcia, R., 2004. Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition. *IEEE Transactions on Geoscience and Remote Sensing* 42 (6), 1291–1299.
- Ballester, C., Caselles, V., Igual, L., Verdera, J., 2006. A variational model for P+XS image fusion. *International Journal of Computer Vision* 69 (1), 43–58.
- Bamberger, R.H., Smith, M.J.T., 1992. A filter bank for the directional decomposition of images: theory and design. *IEEE Transaction on Signal Processing* 40 (4), 882–893.
- Choi, M., 2006. A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter. *IEEE Transaction on Geoscience and Remote Sensing* 44 (6), 1672–1682.
- Coello, C.A., Pulido, G.T., Lechuga, M.S., 2004. Handling multiple objectives with particle swarm optimization. *IEEE Transaction on Evolutionary Computing* 8 (3), 256–279.
- Colditz, R., Wehrmann, T., Bachmann, M., Steinnocher, K., Schmidt, M., Strunz, G., Dech, S., 2006. Influence of image fusion approaches on classification accuracy: A case study. *International Journal of Remote Sensing* 27 (15), 3311–3335.
- Da Cunha, A.L., Zhou, J.P., Do, M.N., 2006. The nonsubsampled contourlet transform: Theory, design and applications. *IEEE Transactions on Image Processing* 15 (10), 3089–3101.
- Daugman, J.G., 1980. Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research* 20 (10), 847–856.
- Do, M.N., Vetterli, M., 2005. The contourlet transform: An efficient directional multiresolution image representation. *IEEE Transactions on Image Processing* 14 (12), 2091–2106.
- Do, M.N., Vetterli, M., 2001. Pyramidal directional filter banks and curvelets. In: Proc. IEEE International Conference on Image Processing, ICIP, Thessaloniki, Greece. 7–10 October, 4 p.
- Du, Q., Younan, N.H., King, R., Shah, V.P., 2007. On the performance evaluation of pan-sharpening techniques. *IEEE Geoscience and Remote Sensing Letters* 4 (4), 518–522.
- Ehlers, M., 2004. Spectral Characteristics Preserving Image Fusion Based on Fourier Domain Filtering. In: Proc. of the Remote Sensing for Environmental Monitoring, GIS Applications, and Geology IV, Maspalomas, Spain. 14–16 September. 13 p..
- Fasbender, D., Radoux, J., Bogaert, P., 2008. Bayesian data fusion for adaptable image pansharpening. *IEEE Transaction on Geoscience and Remote Sensing* 46 (6), 1847–1857.
- Garzelli, A., Nencini, F., 2006. Pan-sharpening of very high resolution multispectral images using genetic algorithms. *International Journal of Remote Sensing* 27 (15), 3273–3292.
- Garzelli, A., Nencini, F., Capobianco, L., 2008. Optimal MMSE pan-sharpening of very high resolution multispectral images. *IEEE Transaction on Geoscience and Remote Sensing* 46 (1), 228–236.
- Hui, Y.X., Cheng, J.L., 2008. Fusion algorithm for remote sensing images based on nonsubsampled contourlet transform. *Acta Automatica Sinica* 34 (3), 274–281.
- Ioannidou, S., Karathanassi, V., 2007. Investigation of the dual-tree complex and shift-invariant discrete wavelet transforms on Quickbird image fusion. *IEEE Geoscience and Remote Sensing Letters* 4 (1), 166–170.
- Kennedy, J., Eberhart, R.C., 1997. A discrete binary version of the particle swarm algorithm. In: Proc. of the IEEE Conference on Systems, Man, and Cybernetics. Piscataway, New Jersey. 12–15 October. pp. 4104–4109.
- Kingsbury, N.G., 2001. Complex wavelets for shift invariant analysis and signals. *Applied and Computational Harmonic Analysis* 10 (3), 234–253.
- Li, X., 2003. A non-dominated sorting particle swarm optimizer for multiobjective optimization. In: Proc. of the Genetic and Evolutionary Computation Conference, Chicago, USA. 3 August. pp. 37–48.
- Nguyen, T.T., Oraintara, S., 2008. The shiftable complex directional pyramid—Part 1: Theoretical aspects. *IEEE Transaction on Signal Processing* 56 (10), 4651–4660.
- Park, S.I., Smith, M.J., Mersereau, R.M., 2004. Improved structure of maximally decimated directional filter banks for spatial image analysis. *IEEE Transaction on Image Processing* 13 (11), 1424–1431.
- Pohl, C., Gendersen, J.L., 1998. Multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing* 19 (5), 823–854.
- Ranchin, T., Wald, L., 2000. Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation. *Photogrammetric Engineering and Remote Sensing* 66 (1), 49–61.

- Rockinger, O., 1997. Image Sequence Fusion Using a Shift Invariant Wavelet Transform. In: Proc. IEEE International Conference on Image Processing, ICIP. Santa Barbara, USA. 26–29 October. pp. 288–291.
- Selesnick, I.W., Baraniuk, R.G., Kingsbury, N.G., 2005. The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine* 22 (6), 123–151.
- Sierra, M.R., Coello, C.A., 2006. Multi-Objective Particle Swarm Optimizers: A Survey of the State-of-the-Art. *International Journal of Computational Intelligence Research* 2 (3), 287–308.
- Simoncelli, E.P., Freeman, W.T., Adelson, E.H., Heeger, D.J., 1992. Shiftable multiscale transform. *IEEE Transactions on Information Theory* 38 (2), 587–607.
- Thomas, C., Ranchin, T., Wald, L., Chanussot, J., 2008. Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Transactions on Geoscience and Remote Sensing* 46 (5), 1301–1312.
- Vijay, Sh.P., Younan, N.H., King, R.L., 2008. An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets. *IEEE Transaction on Geoscience and Remote Sensing* 46 (5), 1323–1335.
- Wahlen, J., 2002. Comparison of standard and image-filter fusion techniques. In: Zanasi, A., Brebbia, C.A., Ebecken, N.F.F., Melli, P. (Eds.), *Data Mining III*. WIT Press, Southampton, UK.
- Wald, L., 1999. Some terms of reference in data fusion. *IEEE Transaction on Geoscience and Remote Sensing* 37 (3), 1190–1193.
- Wald, L., 2000. Quality of high resolution synthesized images: Is there a simple criterion?. In: Proc. of the third conference Fusion of Earth data: Merging point measurements, raster maps and remotely sensed images. Sophia Antipolis, France. 26–28 January. 6 p.
- Wang, Z., Bovik, A.C., 2002. A Universal Image Quality Index. *IEEE Signal Processing Letters* 9 (3), 81–84.
- Watson, A.B., 1987. The cortex transform: Rapid computation of simulated neural images. *Computer Vision, Graphics, and Image Processing* 39 (3), 311–327.
- Yuhas, R.H., Goetz, A.F.H., Boardman, J.W., 1992. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm. In: Proc. Summaries of the Third Annual JPL Airborne Geoscience Workshop. 1 Jun. pp. 147–149.
- Zhang, Y., 2004. Understanding image fusion. *Photogrammetric Engineering and Remote Sensing* 70 (6), 657–661.
- Zhang, Y., Hong, G., 2005. An IHS and wavelet integrated approach to improve pan-sharpening visual quality of natural colour IKONOS and QuickBird images. *Information Fusion* 6 (3), 225–234.
- Zhang, Y., Wang, R., 2004. Multi-resolution and multi-spectral image fusion for urban object extraction. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 35 (Part B3), 960–966.
- Zhou, J., Civco, D.L., Silander, J.A., 1998. A wavelet transform method to merge Landsat TM and SPOT panchromatic data. *International Journal of Remote Sensing* 19 (4), 743–757.