ISPRS Journal of Photogrammetry and Remote Sensing, Volume 152, June 2019, Pages 34-48
DOI:10.1016/j.isprsjprs.2019.04.001

Remote sensing image fusion via compressive sensing

Morteza Ghahremani[a,*], Yonghuai Liu[b], Peter Yuen[c] and Ardhendu Behera[b]

[a] Department of Computer Science, Aberystwyth University, Ceredigion, United Kingdom
[b] Department of Computer Science, Edge Hill University, Ormskirk, United Kingdom
[c] Centre for Electronic Warfare, Electro-Optics, Image and Signal Processing Group, Cranfield University, Swindon, United Kingdom

ARTICLE INFO

Keywords:
pan-sharpening
Compressive sensing
multiscale dictionary
panchromatic data
multispectral data

ABSTRACT

In this paper, we propose a compressive sensing-based method to pan-sharpen the low-resolution multispectral (LRM) data, with the help of high-resolution panchromatic (HRP) data. In order to successfully implement the compressive sensing theory in pan-sharpening, two requirements should be satisfied: (*i*) forming a comprehensive dictionary in which the estimated coefficient vectors are sparse; and (*ii*) there is no correlation between the constructed dictionary and the measurement matrix. To fulfill these, we propose two novel strategies. The first is to construct a dictionary that is trained with patches across different image scales. Patches at different scales or equivalently multiscale patches provide texture atoms without requiring any external database or any prior atoms. The redundancy of the dictionary is removed through K-singular value decomposition (K-SVD). Second, we design an iterative $l_1 - l_2$ minimization algorithm based on alternating direction method of multipliers (ADMM) to seek the sparse coefficient vectors. The proposed algorithm stacks missing high-resolution multispectral (HRM) data with the captured LRM data, so that the latter is used as a constraint for the estimation of the former during the process of seeking the representation coefficients. Three datasets are used to test the performance of the proposed method. A comparative study between the proposed method and several state-of-the-art ones shows its effectiveness in dealing with complex structures of remote sensing imagery.

1. Introduction

Remote sensing satellites utilise spectral sensors for capturing the images of the globe in arbitrary spectral and spatial resolutions. Spectral sensors are capable to deliver high spectral resolution imagery, but the spatial resolution of satellite imagery is normally low. To provide spectral sensors with high-spatial resolution capability is very challenging due to the constraints and limitations on the satellite's payload, cost, complexity and bandwidth. Instead of capturing high-spatial resolution multispectral (HRM) imagery directly, an alternative solution is to capture a single high spatial resolution panchromatic (HRP) image through high-spatial resolution imaging sensors. Then, via image fusion techniques known as pan-sharpening in short, the spatial resolution of the low-spatial resolution multispectral (LRM) images can be enhanced by the HRP data [1; 2; 3; 4]. In most remote sensing applications, the HRM images are more interpretative and thus are in greater demand for numerous applications such as for geological research [5], classification [6], object detection [7; 8], segmentation [9], soil moisture [10] etc.

In the literature, pan-sharpening techniques can be broadly classified into four main categories: component substitution-based, multiresolution analysis-based, model-based and superresolution-based methods. Popular component substitution-based methods include the intensity-hue-saturation (IHS) method [11], Gram-Schmidt [12] and principal component analysis [13]. The computational cost of such methods is relatively low. Component substitution-based methods often provide high-quality spatial details but they usually also introduce spectral distortion in the pan-sharpened multispectral data due to the gray scale value difference between the LRM and the HRP images.

Multiresolution analysis-based methods map the LRM and HRP images into the frequency domain, and their low- and high-frequency components are then fused together via appropriate fusion rules. The most powerful multiresolution transforms applied in pan-sharpening have been the decimated wavelets [14], complex wavelets [15], undecimated wavelets [16], curvelets [17; 3], support value transform [18], contourlets [19; 20] and ripplets [21].

*Corresponding author.
mog9@aber.ac.uk/morteza.ghahremani.b@gmail.com (M. Ghahremani); yonghuai.liu@edgehill.ac.uk (Y. Liu); p.yuen@cranfield.ac.uk (P. Yuen); beheraa@edgehill.ac.uk (A. Behera)
ORCID(s):

Published by Elsevier. This is the Author Accepted Manuscript issued with: Creative Commons Attribution Non-Commercial No Derivatives License (CC:BY:NC:ND 4.0).
The final published version (version of record) is available online at DOI:10.1016/j.isprsjprs.2019.04.001. Please refer to any applicable publisher terms of use.

To effectively implement multiresolution analysis-based methods, proper models for extracting the spatial details of the HRP data and how they are then injected into the LRM bands are required (see [22; 23; 24] and the references therein for more details). In recent years, several algorithms and methods based on deep learning have been proposed for pan-sharpening [25; 26; 27]. The computational cost of such methods are high as they require considerable time for learning. Comprehensive surveys on the above categories can be found in recent work [4; 28; 29; 30].

Spectral superresolution has been a popular framework in recent remote sensing image fusion applications such as multispectral superresolution (or, equivalently, LRM pan-sharpening) and hyperspectral superresolution. Multispectral superresolution aims to reconstruct the unknown HRM images from the available LRM and HRP data. The superresolution approach in pan-sharpening can be further categorized into two main groups: the reconstruction-based methods and the compressive sensing-based methods. The former are based on enforcing a reconstruction constraint and imposing posterior knowledge on high-resolution images [31; 32; 33]. The vast amount of research has been conducted based on the compressive sensing theory due to its great potential in recovering signals/images from their compressed ones [34; 35]. The compressive sensing theory has been also widely used in hyper-spectral superresolution [36; 37; 38; 39; 40]. In hyper-spectral superresolution, a set of hyperspectral and LRM images are modeled as compressed measurements of the unknown high-spatial resolution hyperspectral (HRH) images and the unknown HRH images can be recovered under the compressive sensing theory [41].

The compressive sensing theory was first attempted by Li and Yang [42] in the pan-sharpening work and it was then improved by Jiang *et al.* [43]. The results are impressive. Their methods, however, have two main shortcomings. Firstly, the learning process of these methods requires abundant high-resolution HRM samples, which are often unavailable; even if they are available, a large dictionary usually leads to extensive computation. Secondly, although the authors appropriately model the HRP data as a linear combination of the unknown HRM data, the requirement of a predetermined weight vector makes the method not very usable in practise. In fact, applying a fixed global predetermined weight vector is not reliable in the patch-wise analysis. The latter drawback can also be seen in [44; 45; 46]. Li *et al.* [44] applied a ridge regression-like method to recover the HRM dictionary from the available HRP data which makes the results coarse and inconvenient, as the HRP images do not contain valuable spectral information. In [45], the authors employed a component substitution method to train the HRM dictionary which required high computational effort and the results were seen to be biased.

To tackle these problems, the authors in [47] proposed a method that uses a coupled dictionary for the reconstruction of the HRM data. The dictionary obtained through the HRP data is known as the high-resolution (HR) dictionary and, likewise, a low-resolution (LR) dictionary is trained from the spatially degraded HRP image. This method, which is called 'SparseFI' by the authors, presents each patch of the LRM data in the LR dictionary and the same coefficients are then applied to the HR dictionary to reconstruct the HRM image patches. This approach has been modified later by several researchers [48; 49]. SparseFI-like methods often make the fused results smooth as they use just the HRP data to represent each band of the LRM data [24; 50].

Dictionary learning plays an important role in the reconstruction of HRM images. In this study, a new dictionary learning method suitable for the remote sensing image fusion applications is proposed. The proposed dictionary uses patches across different scales as they provide texture atoms without requiring any external database or prior atoms. The redundant patches in the dictionary are then removed by K-singular value decomposition (K-SVD) [51]. Since the dictionary learning utilises patches across different scales, this methodology is called as 'multiscale dictionary'. The proposed multiscale dictionary is the first successful attempt to tackle and to solve the low spatial resolution weakness in the LRM data. It is shown in this work that the proposed dictionary increases the robustness of the estimated representation coefficients substantially. However, in order to successfully implement the compressive sensing theory for remote sensing applications, construction of an overcomplete dictionary is not sufficiently enough. This is because the structures of remote sensing data are so complex and nonlinear that it requires a dedicated algorithm to seek for the sparse representation coefficients. Thus, an iterative algorithm based on alternating direction method of multipliers (ADMM) [52] is also proposed to seek for the sparse coefficients. In the proposed framework, the estimated HRM data is stacked together with the given LRM input data for the estimation of the sparse representation coefficients, which is seen to be more robust to the missing patches. In addition, the LRM data patches had been used as a guide for the estimation of their corresponding HRM ones. Effectively, these strategies have helped to make the HRM data patches very close to their low resolution ones in both the sparse and weight vectors.

The remainder of this paper is organized as follows. The sparse representation and the compressive sensing theories are briefly reviewed in Section 2, Section 3 details the proposed method, and experimental results and discussions are provided in Section 4. Finally, Section 5 draws conclusions and presents possible future work.

2. Compressive sensing and sparse representation

The aim of sparse representation is to approximate a target signal/image using a linear combination of a few elementary observations/measurements. Compressive sensing aims to reconstruct the missing signal/image as accurately as possible with few measurements. As our proposed method is based on sparse representation and compressive sensing [34; 35], they are briefly outlined below for its easy understanding and implementation.

2.1. Sparse representation theory

It is well known that patches of images in particular those of the natural scenery, s, tend to be repetitive regionally and thus they can be readily compressed into distinctive sparse elements. This means that patches s can be represented by a linear combination α of words/atoms of a dictionary $\mathbf{D}$:

$$s = \mathbf{D}\alpha, \tag{1}$$

where 's' is a '$\sqrt{n} \times \sqrt{n}$' patch of the original image; '$\mathbf{D} \in R^{n \times k}$' is an overcomplete dictionary with 'k' atoms, where $n \ll k$, and $\alpha = (\alpha_1, ..., \alpha_k)^{\mathrm{T}}$ is a coefficient vector of size '$k \times 1$'. The underlying motivation for using sparse theory is to project the high-dimensional s into some lower-dimensional subspace, i.e.,

$$\min_{\alpha \in R^k} \| \alpha \|_0 \quad subject\ to \quad s = \mathbf{D}\alpha, \tag{2}$$

where '$\| \, . \, \|_0$' is a pseudo-norm, which counts the number of non-zero elements of α.

2.2. Compressive sensing theory

Compressive sensing theory addresses a specific type of sparse approximation problem. Assume that we observe a small set of measurements (denoted by 'y') of the original image:

$$y = \mathbf{M}s = \mathbf{MD}\alpha, \tag{3}$$

where '$\mathbf{M}$' is a measurement matrix. According to the compressive sensing theory, the sparse α will allow the recovery of the original image from the partial observations/measurements y through the following objective function, provided that the measurement matrix $\mathbf{M}$ is uncorrelated/incoherent with the dictionary $\mathbf{D}$:

$$\min_{\alpha \in R^k} \| \alpha \|_1 \quad subject\ to \quad \| y - \mathbf{MD}\alpha \|_2^2 \leq \epsilon, \tag{4}$$

where 'ϵ' is the error tolerance. As the compressive sensing theory requires certain sparsity conditions to function [34; 35], the optimisation of the pseudo-norm in Eq. (2) is an NP-hard problem, which can be relaxed by replacing it with the convex function l_1-norm. In pan-sharpening, the LRM and HRP data patches can be modeled as the measurement matrix of the to-be recovered HRM. In other words, the goal is to reconstruct the HRM data patches [parameter s in Eq. (3)] from the limited measurements y which are the LRM and HRP data patches in this case. In the next section, the relationship between the measurements and the to-be recovered HRM data will be discussed in more detail.

3. Proposed Method

In this section, the proposed method which implements compressive sensing for image fusion is presented. The proposed method involves three main steps: firstly the development of an appropriate fusion framework dedicated to the pan-sharpening task using compressive sensing principles; Secondly, the development of a multiscale dictionary wherein the estimated coefficient vectors are sparse; and finally, the design of an iterative algorithm for finding the sparse coefficient vectors more efficiently. The details of these 3 steps are given in the following subsections.

3.1. Fusion framework

The objective of pan-sharpening is to make pan-sharpened LRM bands as similar as possible to the ideal HRM bands captured by the narrow-band LRM sensor if it had the same spatial resolution as the HRP band captured [53]. Due to the absence of information in the to-be recovered HRM imagery, the recovery could only be made through the

estimation from the available HRP and LRM data. Conventionally, pan-sharpening adopts fusion framework strategy using a mixture of component substitution and model-based frameworks for solving the inverse problem:

$$M_{\mathrm{HR}}^{(k)} = \bar{\mathrm{M}}_{\mathrm{LR}}^{(k)} + (\mathrm{P} - \bar{I}_{\mathrm{LR}}).\mathbf{G}(k), \qquad k = 1, ..., L. \tag{5}$$

In this equation, 'M_{LR}' and 'P' are the given LRM and HRP input images, respectively; 'I_{LR}' represents the low-spatial resolution intensity (LRI) components to be derived from the LRM images; 'L' and the bar denote the number of LRM bands and the up-sampling operation, respectively; and '$\mathbf{G}$' is a weight matrix that modulates the extent of spatial details to be injected into each LRM band as according to the adopted model [22].

This methodology is relatively straightforward to implement and it also gives high-quality spatially pan-sharpened results. The complexity in this framework is the evaluation of the weighting matrix $\mathbf{G}$ which can be derived locally or globally. However, this model occasionally gives result with spectral/colour distortions especially when there are large radiometric differences between the HRP data and the estimated LRI component patches. One way to mitigate this drawback is to employ the high-spatial resolution intensity (HRI) component instead of the HRP data during the pan-sharpening process [21]:

$$M_{\mathrm{HR}}^{(k)} = \bar{\mathrm{M}}_{\mathrm{LR}}^{(k)} + (I_{\mathrm{HR}} - \bar{I}_{\mathrm{LR}}), \qquad k = 1, ..., L, \tag{6}$$

where 'I_{HR}' denotes the HRI component which can be derived from the HRM data. Just as I_{LR} represents the intensity component of the LRM images, I_{HR} is the intensity component of the HRM images. If the HRI and LRI components are meticulously designed, the term $(I_{\mathrm{HR}} - \bar{I}_{\mathrm{LR}})$ in Eq. (6) exhibits smaller dissimilarity than the term $(\mathrm{P} - \bar{I}_{\mathrm{LR}})$ in Eq. (5). The smaller the dissimilarity in the second term of Eq. (6), the smaller the spectral distortions in the result $M_{\mathrm{HR}}^{(k)}$. However, due to the absence of the HRM data which is to be recovered, the HRI component is possible to be reconstructed in the light of compressive sensing theory. As it will be discussed later, the LRM data has been deployed for the reconstruction of the HRI component to reduce the discrepancy between I_{HR} and I_{LR} in Eq. (6). This is possible when both I_{HR} and I_{LR} components are estimated under the same weight vector ω and the same sparse coefficient vector α, resulting in lower spectral distortion and higher robustness to noise and outliers.

In this paper, the high- and low-spatial resolution intensity components are designated by I_{HR} and I_{LR} respectively and 'HR' and 'LR' are acronyms for the high- and low-spatial resolution respectively. The HR intensity component is assumed to contain 'N' patches given by $\mathbf{X} = \{X^i\}_{i=1}^N$. The size of patches is '$\beta \times \beta$' and they are ordered lexicographically as column vectors, i.e. $X^i = (X^i_{1,1}, ..., X^i_{1,\beta}, ..., X^i_{\beta,1}, ..., X^i_{\beta,\beta})^{\mathrm{T}}$, where 'T' is the transpose operation. Likewise, the LR intensity component patches are given by $\mathbf{x} = \{x^i\}_{i=1}^N = \{(x^i_{1,1}, ..., x^i_{1,B}, ..., x^i_{B,1}, ..., x^i_{B,B})^{\mathrm{T}}\}_{i=1}^N$, where '$B$' is the LR patch size. Depending on the spatial characteristics of the remote sensing imagery, the patch size 'B' can be in the range of {6,7,...,10}. The relationship between the B and β is $\beta = p \times B$, where 'p' is the ratio of the spatial resolution between the LRM and the corresponding HRP images (e.g. 4 for the QuickBird data). The ith patch of the to-be recovered HRM image over 'L' spectral bands is in the form of a matrix $\mathbf{Y}^i = \{Y^{i,(j)}\}_{j=1}^L$, where $Y^{i,(j)}$ is the lexical order of the ith patch in the jth spectral band. Thus, the dimension of the matrix $\mathbf{Y}^i$ is '$\beta^2 \times L$'. Likewise, the ith patches of the LRM images is in the form of a '$B^2 \times L$' matrix $\mathbf{y}^i = \{y^{i,(j)}\}_{j=1}^L$, where $y^{i,(j)}$ is the '$B^2 \times 1$' lexical order of the ith patch of the LRM data in the jth spectral band.

The methodology for the reconstruction of the HRI component that will be utilised for the estimation of the missing HRM data is given here. To this end, the characteristics of the satellite imagery are summarized in Table 1. We first develop three assumptions about the satellite data and then formulate them to estimate the desired HRI and LRI components.

- Assumption I

It is seen from Table 1 that the HRI component and the HRP image share similar spatial and spectral characteristics, i.e. both of them exhibit similar spatial and spectral resolutions. Thus, it can be assumed that:

$$X^i = \mathbf{D}\alpha^i, \qquad i = 1, ..., N, \tag{7}$$

where '$\mathbf{D}$' is the '$\beta^2 \times J$' dictionary learnt from the HRP data. The reconstructed HRI component patches X are sparse in the ith coefficient vector of $\alpha^i = \{\alpha^i_j\}_{j=1}^J$, and '$N$' is the total number of patches in the HRI component. $\mathbf{D}$ and α will be discussed below in Sections 3.2 and 3.3, respectively.

Table 1
The characteristics of the satellite imagery and their corresponding notations used in this study.

Symbol	Definition	Spatial Resolution	Spectral Resolution	Patch-Based Symbol
M_{HR}	HRM (to be synthesized)	High	High	$\mathbf{Y}$
M_{HR}	LRM (available)	Low	High	$\mathbf{y}$
P	HRP (available)	High	Low	-
I_{HR}	HRI (to be synthesized)	High	Low	x
P_{L}	LRP (available)	Low	Low	u
I_{LR}	LRI (to be synthesized)	Low	Low	X

- Assumption II

The LRI component patches $\mathbf{x}$ may plausibly be assumed to be the down-sampled versions of their corresponding HRI ones $\mathbf{X}$, i.e.

$$x^i = \mathbf{M}X^i, \qquad i = 1, ..., N, \tag{8}$$

where '$\mathbf{M}$' is the decimation/degradation matrix or equivalently the measurement matrix of size $B^2 \times \beta^2$ written as $\mathbf{M} = p^{-2} \times (\mathbf{I}_{B\times B} \otimes (\mathbf{1}^T \otimes (\mathbf{I}_{B\times B} \otimes \mathbf{1}^T)))$; '$\mathbf{I}$' in the matrix '$\mathbf{M}$' is an identity matrix, '$\mathbf{1}$' is an '$L \times 1$' vector filled with ones and '$\otimes$' is the Kronecker tensor product.

- Assumption III

Since the HRI component spans over the entire spectral range of the HRM bands, thus, it is plausibly assumed that the HRI component can be approximated by a linear combination of the HRM data:

$$X^i = \mathbf{Y}^i \omega^i, \qquad i = 1, ..., N. \tag{9}$$

In the above equation, 'ω^i' stands for the normalized weight vector of length '$L \times 1$' for the ith patch. Through substituting Eq. (9) into Eq. (7) we have

$$\mathbf{Y}^i_{\beta^2\times L}\,\omega^i_{L\times 1} = \mathbf{D}_{\beta^2\times J}\,\alpha^i_{J\times 1}, \qquad i = 1, ..., N. \tag{10}$$

Accordingly ω^i that leads to the sparse α^i can be searched for each patch i in the above equation ($i \in \{1, ..., N\}$). In other words, under the perspective of the compressive sensing theory, the sparse coefficient vector α^i ensures that the HRI component patch X^i can be accurately recovered from the dictionary $\mathbf{D}$ (see Fig. 1). Likewise, the LRI component patches x^i can be derived via Eqs. (8) and (7) as:

$$x^i = \mathbf{M}X^i = \mathbf{MD}\alpha^i, \qquad i = 1, ..., N. \tag{11}$$

This equation states that the LRI intensity component patch x^i is sparse in the decimated dictionary '$\mathbf{MD}$' (the decimated version of the dictionary $\mathbf{D}$ is denoted by $\mathbf{MD}$). Similar to Eq. (9), the LRI component patch x^i can be obtained from the LRM input data $\mathbf{y}^i$ which can be obtained by inserting Eq. (9) into Eq. (8):

$$x^i = \mathbf{M}\mathbf{Y}^i\omega^i = \mathbf{y}^i\omega^i, \qquad i = 1, ..., N. \tag{12}$$

As mentioned earlier, the LRI component patch x^i can be written as a linear combination of the LRM image patches $\mathbf{y}^i$. The combination of last two equations yields:

$$\mathbf{y}^i_{B^2\times L}\,\omega^i_{L\times 1} = \mathbf{M}_{B^2\times\beta^2}\mathbf{D}_{\beta^2\times J}\,\alpha^i_{J\times 1}, \qquad i = 1, ..., N. \tag{13}$$

This equation is similar to Eq. (10) but in this case it is in the LR space. Equations (10) and (13) play a crucial role in the proposed fusion framework as they provide sparse coefficient vectors α to feed into Eq. (7) for the estimation

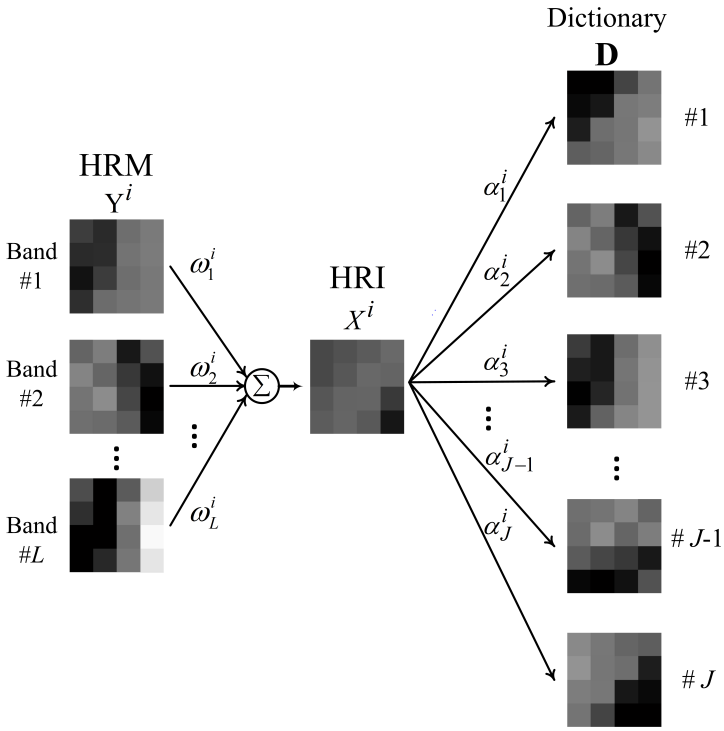


Figure 1: Flowchart of Eq. (10), i.e. $\mathbf{Y}^i\omega^i = \mathbf{D}\alpha^i$; details regarding to the dictionary $\mathbf{D}$ will be discussed in Section 3.2.

of the unknown HRI and LRI component patches. The dimensions of the elements in these equations have been specified to help tracking the fusion process. The flow chart of the overall proposed method is shown in Fig. 2. It depicts how the compressive sensing theory has been implemented in the proposed method. In what follows, we propose a novel 'multiscale dictionary' for pan-sharpening which leads to enhanced sparsity in the coefficients of α. Moreover, a recursive algorithm for seeking the sparse α has been developed. To the best of our knowledge, the proposed multiscale dictionary for pan-sharpening is the first attempt in the remote sensing image fusion research. The proposed pan-sharpening framework fulfils the two fundamental conditions in the compressive sensing theory discussed in the previous section. Firstly, a fixed decimation matrix $\mathbf{M}$ has been deployed as the measurement matrix which exhibits no correlation with the dictionary $\mathbf{D}$. Secondly, it will be shown that α approaches to sparse particularly when the dictionary is complete or in sub-optimal condition.

3.2. Dictionary construction

The strategy of dictionary construction is crucially important for seeking the optimal sparsity of the representation coefficient vector α. During the construction stage, the goal is to select those texture atoms with more faithfully estimated weight vector ω thereby to yield better sparsity in α. In this work, the use of multiple different scales of the HRP data has been deployed for dictionary construction. In other words, the dictionary is obtained through multiscale decomposition of the HRP data thereby it is termed as 'multiscale dictionary'. The multiscale dictionary is established in two steps: (i) the preliminary dictionary construction and (ii) the dictionary refinement. Fig. 3 outlines the flowchart for the formation of the proposed multiscale dictionary. In what follows, the construction of the preliminary dictionary and its advantages are firstly described, then followed by the discussion of the refinement procedure.

Fig. 3 depicts the spatial degradation of the HRP image into 'M' levels with a degradation rate of '$d\%$'. The pyramid is constructed in 'M+1' multiscales in which each has unique characteristic. For instance, if the spatial resolution of the HRP data acquired by IKONOS is $1m$, the spatial resolutions of the 3-level decomposition of the HRP data with a degradation rate of 10% will be $1m$, $1.1m$, $1.2m$ and $1.3m$. This means that the multiscale HRP data is now comprised of the original resolution and the 3 decomposed resolution sets. Each level of the pyramid is then tiled into '$\beta \times \beta$' patches such that the overlap between any two adjacent patches is $c\%$. Each patch is then normalized and centered to

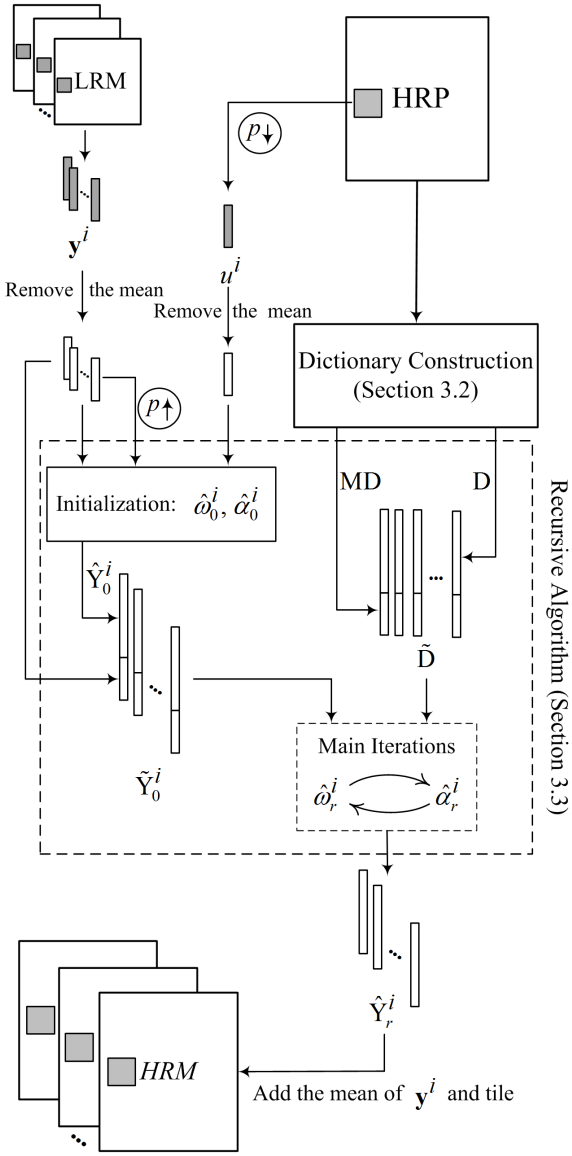


Figure 2: The flowchart of the proposed method.

zero mean. Finally, the pixel values of patches in all levels are then arranged in column vectors to form the preliminary HR dictionary. The parameter $c\%$ depends on the patch size and the spatial characteristic of the HRP data.

There are two major motivations for using the HRP data across different scales in the dictionary construction. Firstly, the LRI component in its original resolution exhibits weak spatial details in comparison to that of the LRP image of the same scene. Secondly, the enrichment of textures of the dictionary atoms has been a key issue in pan-sharpening and this should be considered to enhance the sparsity of the complete dictionary. It is observed that patches of natural imagery tend to be represented by many atoms both within the same scale as well as across different scales. The occurrence of patches across different image scales implicitly provides more appropriate atoms for the patches, thereby the construction of the dictionary can be achieved efficiently even by using a single image without the need of any external database or prior atoms [54].

Fig. 4 shows an example of the sparse coefficient vector of an LRI component patch of the GeoEye-1 data in a five-level multiscale dictionary (i.e. M=0, 1, 2, 3 and 4) with the degradation rate of 12.5%. Note that 'M=0' represents no decomposition of the HRP data. The figure indicates that the number of atoms in the LR dictionary is

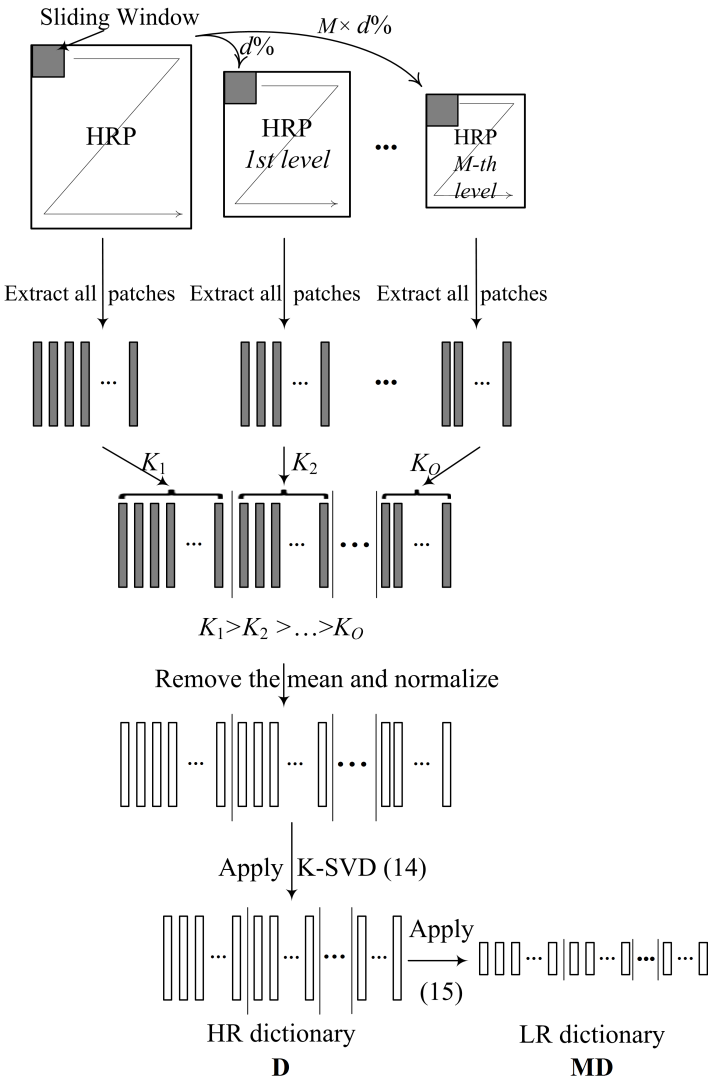
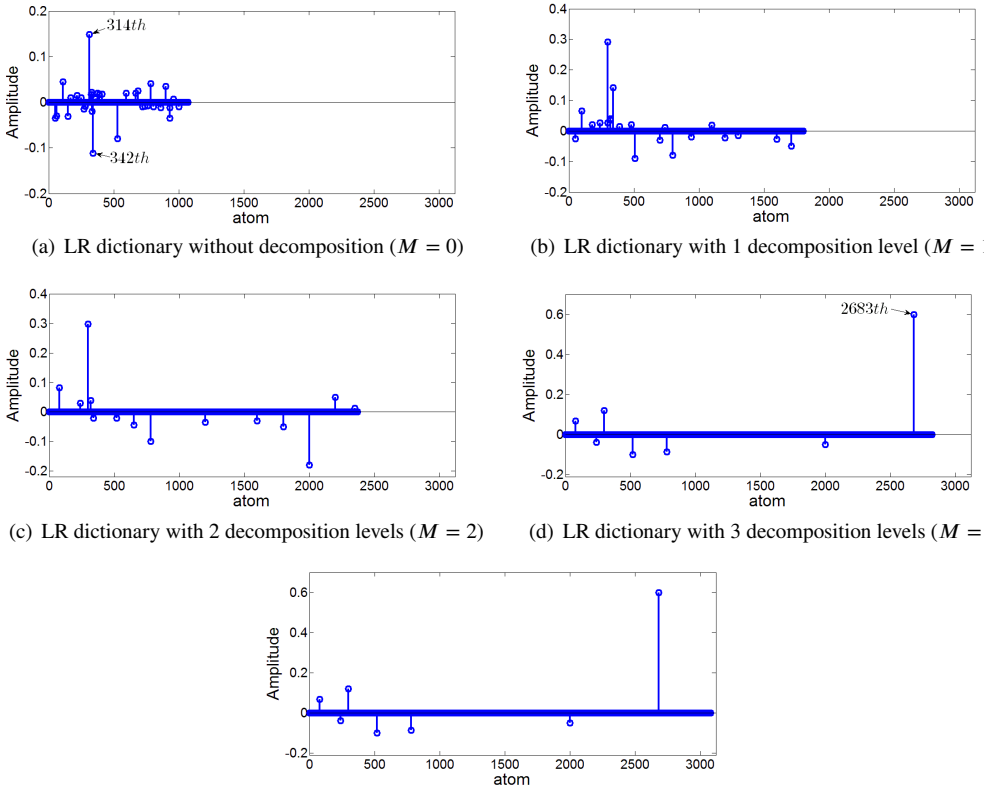


Figure 3: The flowchart for the construction of the proposed multiscale dictionary.

increased when the number of decomposition levels is increased. For example a 3-level decomposition increases the LR dictionary atoms from approximately 1000 at 'M=0' to 2800 at 'M=3'. It is seen that the increased number of atoms tends to substantially increase the sparsity of the dictionary [compare the number of nonzero coefficients in Fig. 4(d) with that in Fig. 4(a)]. This suggests that the more appropriate the atoms in the dictionary, the better the sparsity of the representation coefficient vector. This gives the support to the second claim in this work that states that the sparsity of α is enhanced when multiple scales are used for the construction of the dictionary. Furthermore, it can be observed from Fig. 4 that the most significant atoms in the 'M=0' dictionary, i.e. the 314th and 342th atoms in [Fig. 4(a)], which correspond to the most dominant coefficients of α, are seen to shift towards the end of the vector [2683th atom in Fig. 4(d)] in the 3 decomposition levels (i.e. 'M=3'). This indicates that the LRI component patch is more representative by the 2683th atom in the LR dictionary under the 3 decomposition levels, than that in the LR dictionary with no decomposition level. This gives support of our first claim that the dictionary can be better constructed by using the HRP image in different coarse scales.

Although multiscale decomposition of the HRP data is seen to help enriching the texture atoms of the dictionary, it may also increase the redundancy of the dictionary through recurring itself several times within the same scale as well as across different scales. For instance, the atoms in the fourth decomposition level as shown in Fig. 4(e) are seen



(a) LR dictionary without decomposition ($M = 0$)

(b) LR dictionary with 1 decomposition level ($M = 1$)

(c) LR dictionary with 2 decomposition levels ($M = 2$)

(d) LR dictionary with 3 decomposition levels ($M = 3$)

(e) LR dictionary with 4 decomposition levels ($M = 4$)

Figure 4: The sparse coefficients of a patch of the LRI component of GeoEye-1 presented in a multiscale dictionary with different decomposition levels.

not able to improve the sparsity of vector α significantly. This means that these atoms are redundant without much valuable information and they may even reduce the sparsity of vector α! Redundancy also increases computational load, thereby it has to be penalized. This side effect can be overcome by a dictionary refinement strategy. In this work, the K-singular value decomposition (K-SVD) [51] has been adopted for refining the multiscale dictionary. Dictionary construction aims to seek for the dictionary $\mathbf{D}$ such that it can represent the image data as accurately as possible. Given the HR sample matrix $Z = [Z_1, ..., Z_H]$ (where 'H' is the total number of atoms/samples), the dictionary refinement aims to obtain a dictionary $\mathbf{D}$ such that it satisfies $Z = \mathbf{D}\varphi, \| \varphi \|_0 = K_0$ when neither $\mathbf{D}$ nor φ is known. First, the HR dictionary $\mathbf{D}$ is derived by applying the K-SVD on Z via optimizing the following objective function:

$$\{\hat{\mathbf{D}}, \hat{\phi}\} = \text{argmin}\{\| Z - \mathbf{D}\phi \|_F^2\}, \quad \| \phi \|_0 \leq K_0, \tag{14}$$

where 'K_0' is the sparsity constraint parameter, '$\phi = [\varphi_1, ..., \varphi_J]$' is the sparse representation matrix ($J < H$), and '$\| \, . \, \|_F$ ' is the Frobenius norm of the matrix. To establish the correspondence between the learnt HR and LR atoms, the LR dictionary ($\mathbf{MD}$) is constructed by minimizing the approximation error of $\mathbf{M}Z$ with the same sparse representation coefficients $\hat{\phi}$ in Eq. (14), i.e.

$$\{\hat{\mathbf{MD}}, \hat{\phi}\} = \arg\min_{\mathbf{MD}} \{\| \mathbf{M}Z - \mathbf{MD}\hat{\phi} \|_F^2\} = \mathbf{M}Z\hat{\phi}^T(\hat{\phi}\hat{\phi}^T)^{-1}. \tag{15}$$

Eqs. (14) and (15) not only remove redundant atoms in the HR and LR dictionaries, but also reduce the size of the dictionary that leads to saving computational time substantially. In the following section, the estimation of the coefficient vectors is presented after the multiscale dictionary is constructed.

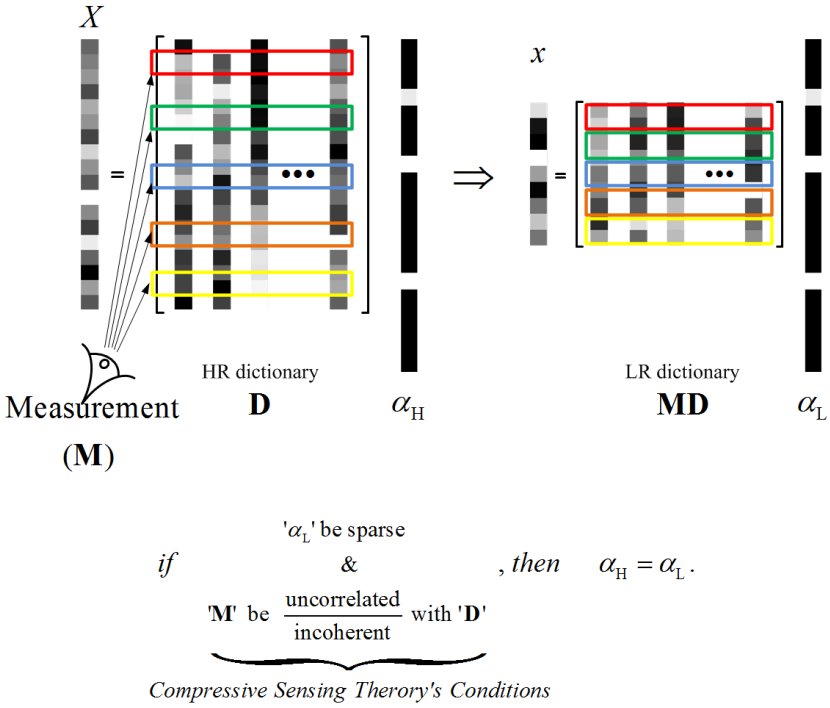


Figure 5: The procedure for the implementation of the compressive sensing theory in the proposed pan-sharpening framework.

3.3. Iterative algorithm for estimating ω and α from Eqs. (10) and (13)

This section presents how Eqs. (10) and (13) can be solved for achieving the desired ω and sparse α. The combination of these equations gives:

$$\min_{\alpha^i,\omega^i} \parallel \alpha^i \parallel_1 \ \& \ \parallel \omega^i \parallel_2 \quad subject\,to \quad \mathbf{Y}^i\omega^i = \mathbf{D}\alpha^i \ \& \ \mathbf{y}^i\omega^i = \mathbf{MD}\alpha^i, \ \ i = 1, ..., N. \tag{16}$$

This equation is valid when the unknown HRI component and the LRI component share the same distribution. Eq. (16) can be rewritten as

$$\min_{\alpha^i,\omega^i} \parallel \alpha^i \parallel_1 \ \& \ \parallel \omega^i \parallel_2 \ subject\,to \ \ \tilde{\mathbf{Y}}^i\omega^i = \tilde{\mathbf{D}}\alpha^i, \ \ i = 1, ..., N. \tag{17}$$

where

$$\tilde{\mathbf{Y}}^i = \begin{bmatrix} \mathbf{Y}^i \\ \mathbf{y}^i \end{bmatrix}, \qquad \tilde{\mathbf{D}} = \begin{bmatrix} \mathbf{D} \\ \mathbf{MD} \end{bmatrix}. \tag{18}$$

Note that Eq. (18) is the concatenation of the LRM patches with its corresponding to-be recovered HRM ones, which may effectively represent sparse coefficient vectors in the concatenated data space. However, it is not necessarily sparse in the LRM space nor in the unknown HRM data individually. Henceforth, the question of the same α in LR whether it also can be used for HR is needed to answer. Due to the unknown nature of the to-be recovered HRM data, it is difficult if not impossible, to assess the sparsity of α in HR. However, if α is sparse in LR then under the prospective of the compressive sensing theory, it may claim that α is also sparse in HR. So it is plausible intuitively to apply α in LR to the HR data (see Fig. 5). It follows that when both the HRI and LRI components are functions of the same ω and α, therefore, the unknown HRM patches can be readily estimated.

An iterative algorithm that recursively optimizes Eq. (17) through an alternative optimization of α, and then ω while holding the other parameter fixed, has been developed for the estimation of the HRM patches by making use of the K-SVD principle. The above algorithm comprises of two optimization problems: an l_1-regularized least squares

problem which is equivalent to the least absolute shrinkage and selection operator (also called Lasso) [55] and an l_2-constrained least square problem called ridge regression that can be solved via ADMM [52]. [52] provides example solutions for the optimization of these two concurrent problems. If the solutions of the Lasso and ridge regression problems obtained by using the ADMM method can be named as 'Lasso-ADMM' and 'ridge-regression-ADMM' respectively, then an approximated α can be estimated via Lasso-ADMM with a fixed ω, and it is then iterated next by fixing the approximated α and subsequently to seek for ω via ridge-regression-ADMM. The process is repeated until the objective function is converged. Note that the recursive algorithm is needed to be initialized via an initial ω due to the non-existence HRM patch. To this end, it is assumed that the ith patch of the LRI component x^i corresponds to that of the LRP image u^i and the initial ω is then estimated. This assumption is reasonable due to the large similarity between the LRI component and the LRP image (see Table 1). The sparse α can be obtained through the recursive updates of α and ω. The details of the proposed recursive algorithm is presented in Algorithm 1 below. It outputs the unknown HRM data patches and subsequently the HRM images after rearranging the patches.

Algorithm 1: Proposed recursive algorithm for solving Eq. (17)

Inputs:

- $\bar{\mathbf{y}}$- upsampled LRM image patches,
- u- LRP image patches,
- $\tilde{\mathbf{D}}$- dictionary,
- α- coefficient vector,
- ω- weight vector,
- τ- termination threshold and
- r- recursive numerator.

For each patch i, $i \in \{1, ..., N\}$, do:
Initialization: Initialize $r = 0$; set $\hat{\mathbf{Y}}^i_{r=0} = \bar{\mathbf{y}}^i$, and solve the following via ridge-regression-ADMM and Lasso-ADMM:
$\hat{\omega}^i_0 = arg\min_{\omega^i_0} \parallel \mathbf{y}^i\omega^i_0 - u^i \parallel_2 + \parallel \omega^i_0 \parallel_2$
and
$\hat{\alpha}^i_0 = arg\min_{\alpha^i_0} \parallel \mathbf{y}^i\hat{\omega}^i_0 - \mathbf{MD}\alpha^i_0 \parallel_2 + \parallel \alpha^i_0 \parallel_1$. Now we can form the initial HRI and LRI components' patches via Eqs. (7) & (8), and subsequently form the initial missing HRM patch: $\hat{\mathbf{Y}}^i_{r=0} = \bar{\mathbf{y}}^i + (\mathbf{D}\hat{\alpha}^i_0 - \overline{\mathbf{MD}\hat{\alpha}^i_0})$. Likewise, we form $\tilde{\mathbf{Y}}^i_0 = \begin{bmatrix} \hat{\mathbf{Y}}^i_0 \\ \mathbf{y}^i \end{bmatrix}$. In the following, we refine the initial HRM patch.
Main Iterations: Increment r by 1 and apply the following steps:

1. **Estimation of** ω^i: Use the ridge-regression-ADMM algorithm to approximate the solution of l_2- constrained least square problem:
   $\hat{\omega^i}_r = arg\min_{\omega^i_r} \parallel \tilde{\mathbf{Y}}^i_{r-1}\omega^i_r - \tilde{\mathbf{D}}\hat{\alpha}^i_{r-1} \parallel_2 + \parallel \omega^i_r \parallel_2$.
2. **Sparse coding stage**: Now seek α^i using Lasso-ADMM:
   $\hat{\alpha}^i_r = arg\min_{\alpha^i_r} \parallel \tilde{\mathbf{Y}}^i_{r-1}\hat{\omega}^i_r - \tilde{\mathbf{D}}\alpha^i_r \parallel_2 + \parallel \alpha^i_r \parallel_1$.
3. **HRM data update**: update the HRM data patches via $\hat{\mathbf{Y}}^i_r = \bar{\mathbf{y}}^i + (\mathbf{D}\hat{\alpha}^i_r - \overline{\mathbf{MD}\hat{\alpha}^i_r})$ and subsequently update $\tilde{\mathbf{Y}}^i_r = \begin{bmatrix} \hat{\mathbf{Y}}^i_r \\ \mathbf{y}^i \end{bmatrix}$.
4. **Stopping update**: If $\parallel \hat{\mathbf{Y}}^i_r - \hat{\mathbf{Y}}^i_{r-1} \parallel_F$ is smaller than a predetermined threshold τ, stop. Otherwise, go to the next iteration.

Output: The desired results are the HRM data patches $\hat{\mathbf{Y}}^i_r$, $i \in \{1, ..., N\}$.

Table 2
The characteristics of the satellite imagery used in this study.

Features	DEIMOS-2		GeoEye-1		QuickBird-2	
	HRP	LRM	HRP	LRM	HRP	LRM
Spatial Resolution (m)	1	4	0.41	1.64	0.61-0.72	2.44-2.88
Spectral Range (nm)	420-720	B*: 420-510	450-800	B: 450-510	760-850	B: 450-520
		G: 510-580		G: 510-580		G: 520-600
		R: 600-720		R: 655-690		R: 630-690
		NIR: 760-890		NIR: 780-920		NIR: 760-900
Dynamic Range	10 bits per pixel per channel		11 bits per pixel per channel		16 bits per pixel per channel	
Location	Vancouver, Canada		Hobart, Australia		Yala, India	

*B: Blue, G: Green, R: Red, NIR: Near InfraRed

4. Experimental Results

Three datasets collected by the Deimos-2[1], GeoEye-1[2] and QuickBird-2[3] satellites imagery have been used to evaluate the proposed method in this study. The scale ratio p is 4 for all the datasets and the size of all the images is 2048 by 2048 pixels. The details of the datasets have been summarized in Table 2.

The fusion results were evaluated by using the widely-popular Wald's synthesis protocol [53]. According to this protocol, the original LRM and the HRP images are preliminarily decimated by the scale ratio p. Then, pan-sharpening is performed on the degraded data and the sharpened results are compared with the original LRM image which plays the role of a ground-truth reference. The pan-sharpened results were assessed by root mean square error (RMSE) as a distortion measurement, cross-correlation coefficient (CC) and a quaternion-based coefficient (Q4) [56] as quality indices. The quality of the fusion results was also calculated by QNR index [57] which is composed of a spectral distortion ($\mathbf{D}_\lambda$) index and a spatial distortion ($\mathbf{D}_s$) index, without requiring the high resolution reference LRM image.

The proposed method was compared with the modified IHS method [58] denoted by MIHS, the two-step sparse coding method [48] denoted by TSSC, the sparse fusion of images method [47] denoted by SparseFI, the local autoregressive model-based method [46] denoted by AM, the blind model-based method [59; 60] denoted by BM[4], and the context-adaptive pan-sharpening method based on segmentation [61] denoted by CAS[5]. The number of multiscale decomposition levels M was set to 3 with a degradation rate of 10% which will be further discussed in the computational time section. In order to maintain a balanced trade-off between performance and computation time in the proposed method, a patch size of 8*8 with an overlap rate of 12.5% and a termination threshold 'τ=0.05' were chosen. All the other competing superresolution-based methods were implemented with the same patch size of 8*8 for fair comparisons. The overlapping area size was set as 3*3 for the TSSC and SparseFI methods, and 2*2 for the AM method, to achieve their optimum performances in terms of quality and computational time. Otherwise all the other competing methods were run using their default parameters.

4.1. Experiments with the Deimos-2 data

Fig. 6(a) shows an example of the LRM image in true color with a spatial resolution of $4m$ captured by the Deimos-2 satellite. The fusion results by different methods are presented in Fig. 6(b)-(h). Magnified subscenes of the LRM and fused results indicated by red arrows have been provided in Fig. 7 to better appraise the pan-sharpening results. Visual comparisons between the pan-sharpened images with respect to the ground-truth reference show that all the methods exhibit good performances on spectral preservation except MIHS and CAS. From the spatial perspective, AM, SparseFI and BM introduce some weak edges. For instance, the edges of the buildings are blurred for the results obtained by SparseFI [Fig. 7(c)], AM [Fig. 7(d)] and BM [Fig. 7(g)], respectively. It is seen that the proposed method

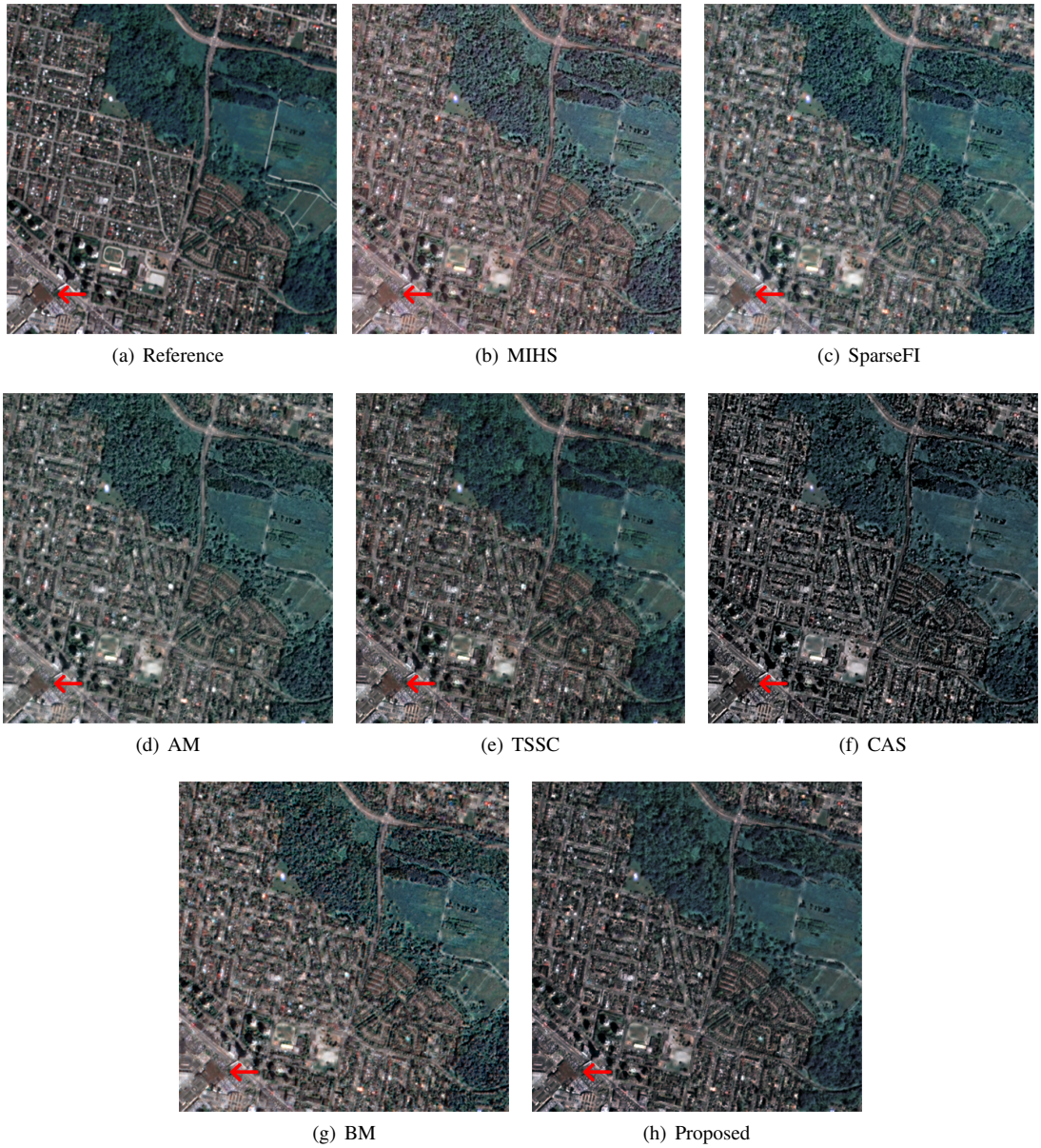
[1]The database is available from http://www.grss-ieee.org/community/technical-committees/data-fusion/2016-ieee-grss-data-fusion-contest/.

[2]The database is available from https://www.satimagingcorp.com/gallery/geoeye-1/.

[3]The database is available from http://glcf.umd.edu/data/quickbird/.

[4]The code is available from https://github.com/qw245/BlindFuse/.

[5]The code is available from http://openremotesensing.net/knowledgebase/context-adaptive-pansharpening-based-on-image-segmentation/.



(a) Reference (b) MIHS (c) SparseFI

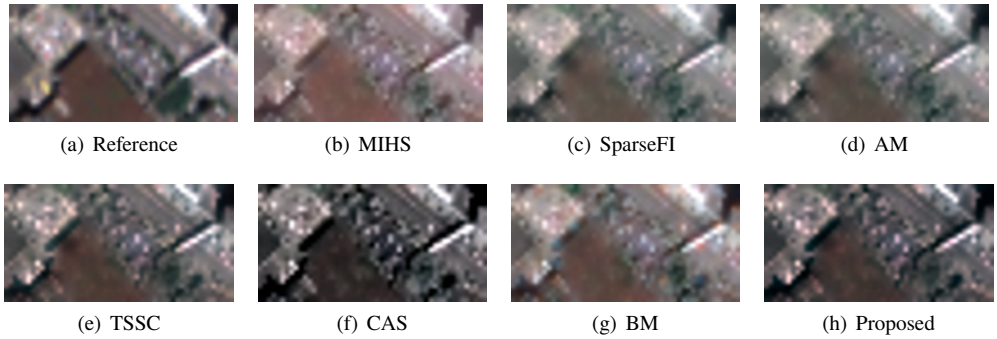
(d) AM (e) TSSC (f) CAS

(g) BM (h) Proposed

Figure 6: Sub-scenes of the Deimos-2 image and the pan-sharpening results (true color composition: RGB).

as shown in Fig. 7(h) provides the best spatial and spectral qualities over all the other competing methods.

Table 3 tabulates quantitative assessment indexes for the Deimos-2 data at reduced resolutions, i.e. by merging $4m$ HRP data with $16m$ LRM data to obtain $4m$ fused images which are then compared with the original $4m$ LRM bands. The best results are shown in bold and the second-best ones are underlined. The quantitative assessment results are consistent with the visual comparison. For all the four bands, the proposed approach provides the best CC and RMSE values. From the Q4 perspective, it also gains the best score, which impels a higher similarity between the pan-sharpened result by the proposed method with respect to the reference ground truth image. In order to calculate the QNR index, the $1m$ HRP image is merged with the $4m$ LRM image to obtain the $1m$ HRM one. The quantitative results are tabulated in Table 4, which indicates that the proposed method has provided the best result in terms of the spatial distortion, spectral distortion and QNR values.



(a) Reference (b) MIHS (c) SparseFI (d) AM

(e) TSSC (f) CAS (g) BM (h) Proposed

Figure 7: Magnified regions indicated by the red arrows in Fig. 6.

Table 3
The quantitative assessment results of different methods with the Deimos-2 data ($4m$ HRP and $16m$ LRM).

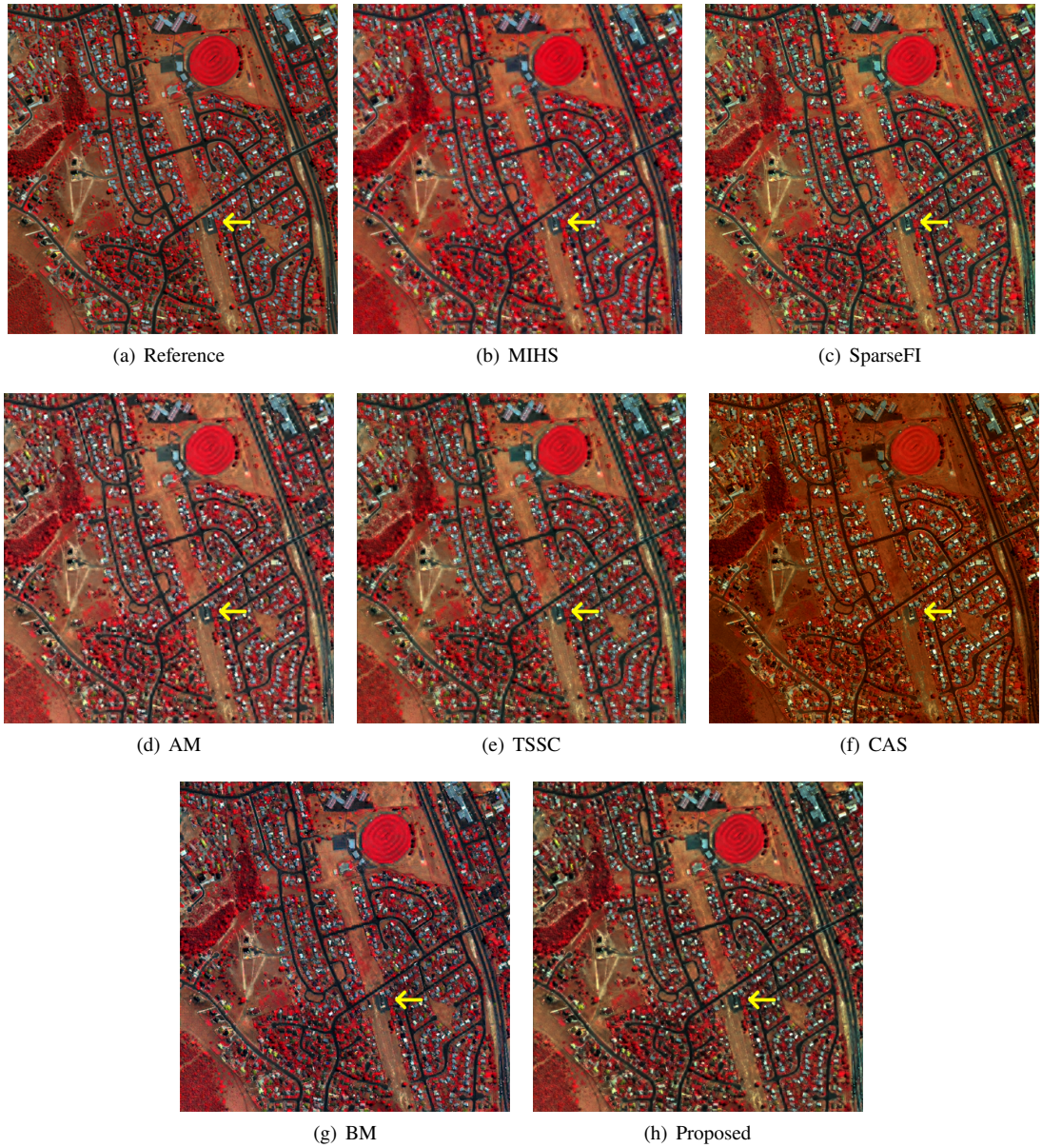
Metrics	Reference Value	MIHS [58]	SparseFI [47]	AM [46]	TSSC [48]	CAS [61]	BM [59]	Proposed
CC_R	1	0.894	0.931	0.923	0.935	0.835	0.883	**0.951**
CC_G	1	0.850	0.928	0.921	0.921	0.847	0.879	**0.949**
CC_B	1	0.854	0.914	0.918	0.927	0.823	0.868	**0.942**
CC_{NIR}	1	0.871	0.921	0.925	0.925	0.895	0.915	**0.941**
$CC_{avg.}$	1	0.867	0.924	0.922	0.927	0.850	0.881	**0.946**
$RMSE_R$	0	16.15	12.85	12.39	12.35	17.23	14.90	**12.02**
$RMSE_G$	0	15.40	13.33	12.52	12.71	16.35	15.82	**12.18**
$RMSE_B$	0	13.56	13.45	12.80	11.81	14.87	14.33	**10.98**
$RMSE_{NIR}$	0	11.32	8.45	8.05	8.89	10.11	10.97	**7.84**
$RMSE_{avg.}$	0	14.11	11.64	11.44	11.24	14.64	14.01	**10.78**
$Q4$	1	0.774	0.901	0.906	0.903	0.757	0.792	**0.918**

Table 4
The QNR indexes of different methods with the Deimos-2 data ($1m$ HRP and $4m$ LRM).

Metrics	Reference Value	MIHS [58]	SparseFI [47]	AM [46]	TSSC [48]	CAS [61]	BM [59]	Proposed
$\mathbf{D_s}$	0	0.058	0.068	0.061	0.055	0.095	0.065	**0.043**
$\mathbf{D}_\lambda$	0	0.112	0.063	0.069	0.064	0.174	0.071	**0.057**
QNR (avg.)	1	0.836	0.873	0.874	0.884	0.748	0.869	**0.902**

4.2. Experiments with the GeoEye-1 data

The original LRM image (false color composition: NIR-R-B) with a spatial resolution of $1.64m$ acquired by the GeoEye-1 satellite is shown in Fig. 8(a). The fusion results of the seven different methods are shown in Fig. 8(b)-(h). To meticulously evaluate the performances of various pan-sharpening methods, the magnified subscenes of the LRM and fused results indicated by yellow arrows in Fig. 8 are shown in Fig. 9. Spectral comparisons of the pan-sharpened results with respect to the reference image [Fig. 9(a)], CAS [Fig. 9(f)] is seen to preserve low color information. While MIHS and TSSC have shown some spectral distortions but they are not larger than that of the BM's result. For example, the color of the regions in the reference image is in blood-red while it is in cherry red for the results given by the MIHS [Fig. 9(b)] and TSSC [9(e)] methods respectively. Other methods preserve spectral information especially in the blood-red regions as well as asphalt regions better [Fig. 9(c), (d), (h)]. From the spatial injection perspective, the BM, AM and the proposed methods are more successful than the others; while the edges of the buildings, asphalt road and white region in the BM method, as well as in our proposed method, are sharper and clearer than those produced by



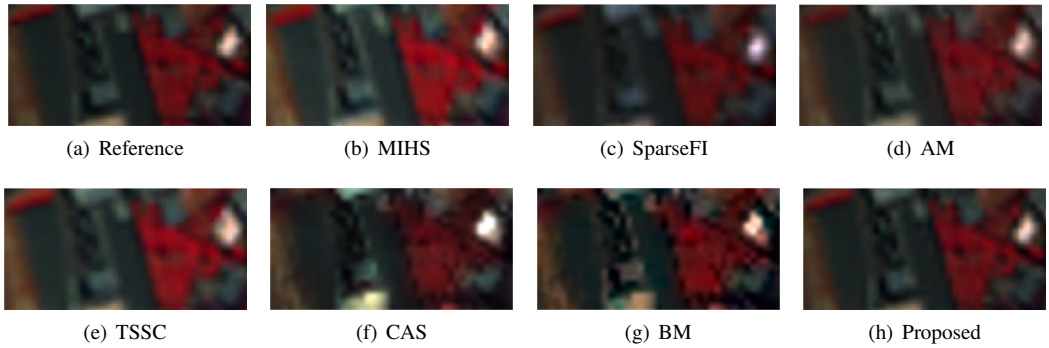
(a) Reference (b) MIHS (c) SparseFI

(d) AM (e) TSSC (f) CAS

(g) BM (h) Proposed

Figure 8: Sub-scenes of the GeoEye-1 image and the pan-sharpening results (false color composition: NIR-R-B).

all the other methods. Also the high spatial injection in the BM method has caused some spectral distortions. Overall, Fig. 9 verifies that the result of the proposed method is spectrally and spatially more similar to the reference image than all the other methods utilised in this study. Tables 5 and 6 tabulate the goodness assessments of pan-sharpened results produced by various methods for the degraded and full scale datasets respectively. From these tables, it can be seen again that the proposed method has achieved the best performance in most of the metrics utilised for the goodness assessments. As mentioned earlier, the spatial information of BM is much higher than the other methods and for this reason it usually yields the smallest spatial distortion.



(a) Reference (b) MIHS (c) SparseFI (d) AM

(e) TSSC (f) CAS (g) BM (h) Proposed

Figure 9: Magnified regions indicated by the yellow arrows in Fig. 8.

Table 5
The quantitative assessment results of different methods with the GeoEye-1 data (1.64m HRP and 6.56m LRM).

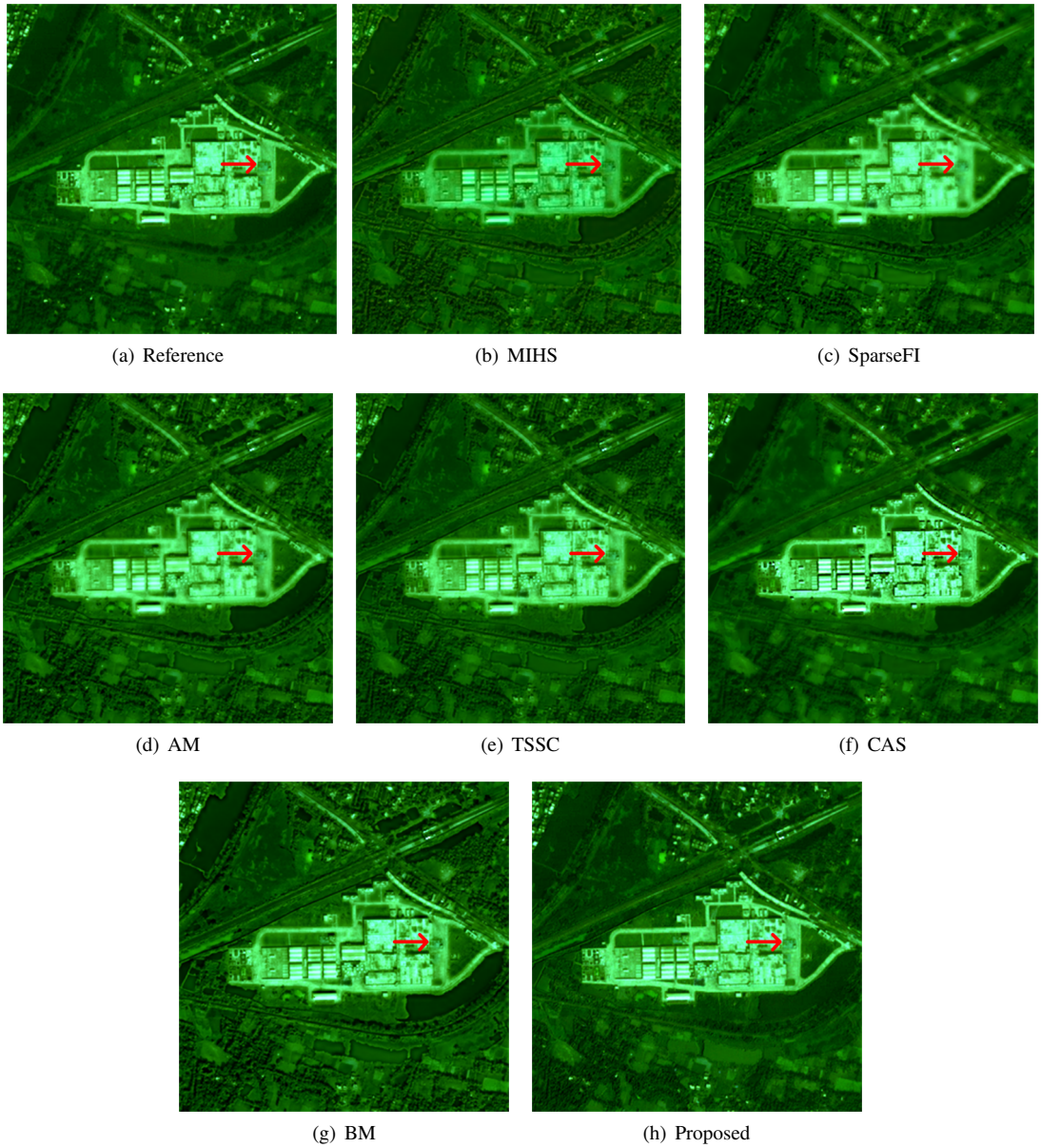
Metrics	Reference Value	MIHS [58]	SparseFI [47]	AM [46]	TSSC [48]	CAS [61]	BM [59]	Proposed
CC_R	1	0.885	0.932	0.947	0.940	0.846	0.826	**0.955**
CC_G	1	0.881	0.934	0.941	0.942	0.842	0.872	**0.957**
CC_B	1	0.875	0.924	0.938	0.936	0.835	0.871	**0.956**
CC_{NIR}	1	0.888	0.938	0.943	0.945	0.811	0.834	**0.966**
$CC_{avg.}$	1	0.882	0.933	0.942	0.939	0.834	0.851	**0.958**
$RMSE_R$	0	15.01	14.52	14.82	15.35	14.85	15.80	**14.25**
$RMSE_G$	0	12.63	12.44	12.55	12.35	13.39	13.01	**11.62**
$RMSE_B$	0	13.54	12.95	12.41	12.32	13.93	13.66	**10.98**
$RMSE_{NIR}$	0	11.80	11.96	11.37	11.18	12.57	11.22	**10.86**
$RMSE_{avg.}$	0	13.25	12.97	12.79	12.81	13.69	13.42	**11.92**
$Q4$	1	0.844	0.909	0.916	0.921	0.896	0.863	**0.944**

Table 6
The QNR indexes of different methods with the GeoEye-1 data (0.41m HRP and 1.64m LRM).

Metrics	Reference Value	MIHS [58]	SparseFI [47]	AM [46]	TSSC [48]	CAS [61]	BM [59]	Proposed
$\mathbf{D_s}$	0	0.115	0.106	0.088	0.113	0.123	**0.068**	0.075
$\mathbf{D}_\lambda$	0	0.148	0.124	0.121	0.118	0.172	0.153	**0.098**
QNR (avg.)	1	0.754	0.783	0.802	0.782	0.726	0.789	**0.834**

4.3. Experiments with the QuickBird-2 data

The fusion results of the QuickBird-2 data produced by different methods are depicted in Fig. 10. Like previous subsections, the magnified sub-scenes of the fused results generated by all the methods are shown in Fig. 11. For this dataset, the pan-sharpened results by BM, CAS, and the proposed methods, shown in Figs. 11(f), 11(g) and 11(h) respectively, exhibit higher degrees of similarity with respect to the reference ground truth [Fig. 11(a)] both spatially and spectrally. The result of SparseFI appears to be spatially smooth. CAS shows high spatial injection in the textured regions but low spatial injection in texture-free regions. BM shows high injection of spatial information in the pan-sharpened result while the proposed method provides a good balance between the spatial and spectral information. Tables 7 and 8 summarise the goodness assessments of these methods for this dataset and the quantitative results are seen to be consistent with the pan-sharpened images presented in the figures. In summary, the performance of the



(a) Reference (b) MIHS (c) SparseFI

(d) AM (e) TSSC (f) CAS

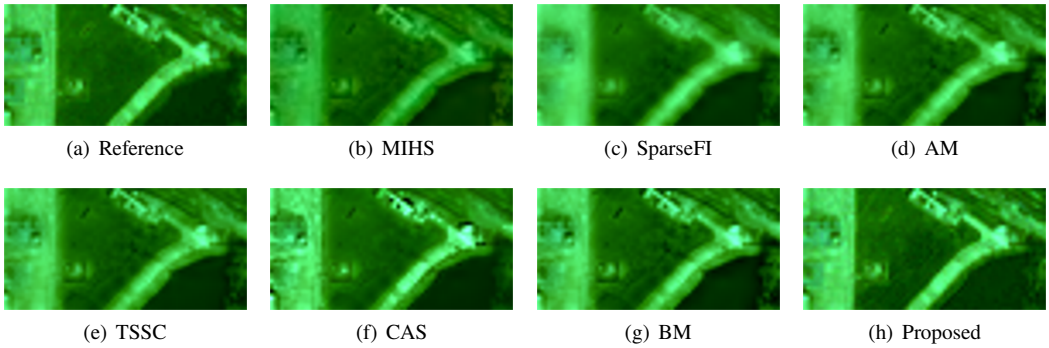
(g) BM (h) Proposed

Figure 10: Sub-scenes of the QuickBird-2 image and the pan-sharpening results (true color composition: RGB).

proposed method again gives the best results in most metrics over all the other competing pan-sharpening algorithms.

4.4. Time cost

All the experiments conducted in this work was performed in MATLAB 2013a using a computer with CPU Core i5/3.20 GHz and 4 GB RAM. The average computational time of different methods for executing the pan-sharpening over the 3 datasets is reported in Table 9. The time cost of the proposed method is comparable to those of TSSC and SparseFI but it is high computationally in comparison to that of the MIHS method, which takes less than one second for performing the task. It is worth mentioning that the proposed method is a recursive algorithm which is not usually favourable from the computational load perspective; however, the time cost is not the primary concern in the present study as it can be markedly reduced via multithreaded GPU implementation.



(a) Reference (b) MIHS (c) SparseFI (d) AM

(e) TSSC (f) CAS (g) BM (h) Proposed

Figure 11: Magnified regions indicated by the red arrows in Fig. 10.

Table 7
The quantitative assessment results of different methods with the QuickBird-2 data (2.44-2.88m HRP and 9.76-11.52m LRM).

Metrics	Reference Value	MIHS [58]	SparseFI [47]	AM [46]	TSSC [48]	CAS [61]	BM [59]	Proposed
CC_R	1	0.794	0.916	0.886	0.897	0.915	0.877	**0.950**
CC_G	1	0.908	0.903	0.925	0.910	0.922	0.900	**0.941**
CC_B	1	0.896	0.914	0.889	0.904	0.917	0.889	**0.938**
CC_{NIR}	1	0.890	0.920	0.924	0.917	0.923	**0.925**	0.921
$CC_{avg.}$	1	0.872	0.913	0.906	0.907	0.919	0.898	**0.938**
$RMSE_R$	0	12.88	10.21	11.78	11.19	10.91	12.43	**9.44**
$RMSE_G$	0	12.05	10.14	11.39	10.18	10.78	11.58	**9.54**
$RMSE_B$	0	12.96	11.09	11.35	11.23	10.45	11.30	**9.82**
$RMSE_{NIR}$	0	13.35	**10.78**	12.26	12.86	11.13	12.02	11.54
$RMSE_{avg.}$	0	12.81	10.56	11.69	11.36	10.81	11.83	**10.09**
$Q4$	1	0.657	0.695	0.689	0.718	0.728	0.704	**0.817**

Table 8
The QNR indexes of different methods with the QuickBird-2 data (0.61-0.72m HRP and 2.44-2.88m LRM).

Metrics	Reference Value	MIHS [58]	SparseFI [47]	AM [46]	TSSC [48]	CAS [61]	BM [59]	Proposed
$\mathbf{D_s}$	0	0.187	0.141	0.149	0.145	0.102	0.176	**0.083**
$\mathbf{D}_\lambda$	0	0.136	0.068	0.089	0.091	0.051	0.120	**0.041**
QNR (avg.)	1	0.702	0.801	0.775	0.773	0.852	0.725	**0.879**

4.5. The number of decomposition levels for constructing the dictionary

This section discusses the relationship between the number of decomposition levels and the sparsity of the dictionary. The effect for using different numbers of decomposition levels in the GeoEye-1 dataset will be discussed here and this can be extended to the other datasets. Figs. 12 and 13 show the quality of the fusion results by RMSE and CC respectively as functions of decomposition levels. It can be seen from these figures that increasing the number of atoms improves the quality of fusion results in general. The improvements are more significant when the first decomposition level is applied and the enhancement seems to be leveled off after 3 decomposition levels. It is observed that the increased number of atoms after 4 decomposition levels could not significantly improve the expressiveness of the dictionary. This is particularly valid for the proposed method: the convergence of the proposed recursive algorithm almost remains the same after 3 decomposition levels. This can be seen from the averaged number of recursions result

Table 9
The average time cost in seconds of different methods over the three datasets.

MIHS [58]	SparseFI [47]	AM [46]	TSSC [48]	CAS [61]	BM [59]	Proposed
0.6	35.3	858	48.7	2.8	7.3	40.2

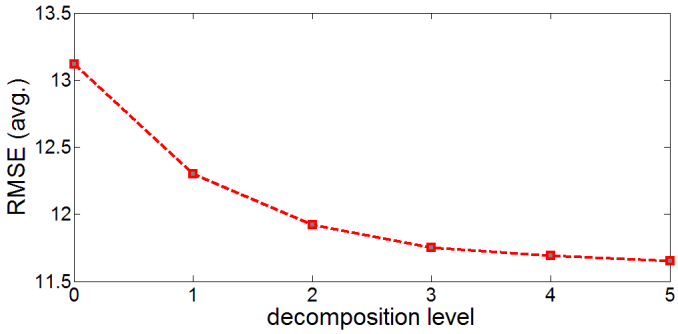


Figure 12: The average root mean square error of the proposed method for different decomposition levels.

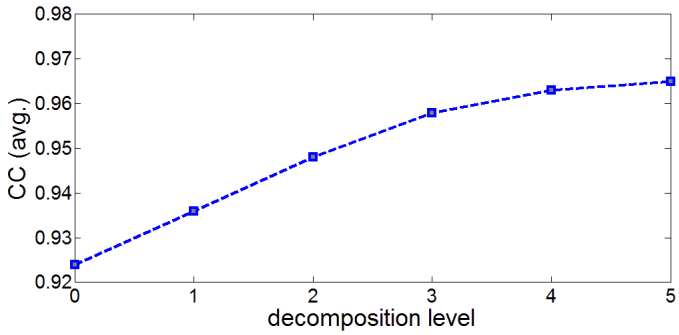


Figure 13: The average CC of the proposed method for different decomposition levels.

as shown in Fig. 14 [it is denoted by r (avg.)]. As noted above, the nonlinear behavior of remote sensing data does not allow the algorithm to seek for better atoms after 'M=3'. The performance of the algorithm is saturated after 3 levels of decomposition. Henceforth, 3 decomposition levels are sufficient for the construction of the proposed multiscale dictionary if a good balance between the run-time and the sparsity is maintained. For this reason 'M=3' has been chosen throughout this paper.

5. Conclusion and Future Work

In this paper, we have proposed a novel pan-sharpening method based on the compressive sensing theory and the dictionary reconstruction through a multiscale decomposition methodology. The novelty in this paper is in the direction of successfully implementing the compressive sensing theory for remote sensing image fusion. The performance of the proposed method has been examined over several datasets. The results have visually and quantitatively verified to demonstrate the effectiveness of the proposed method. For large datasets, the data can be partitioned into several sub-images with small overlap and pan-sharpening can then be applied to each of them independently. As the proposed method is designed for pan-sharpening multispectral data, application of the proposed method to higher dimensions of spectral data like hyper-spectral imagery is possible which will be a good subject for future investigation.

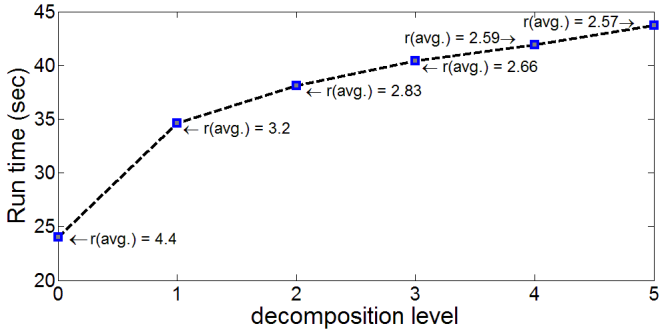


Figure 14: The time cost in seconds of the proposed method for different decomposition levels; the average number of recursions at each scale is denoted by 'r(avg.)'.

Acknowledgments

This work was supported in part by DSTL (DSTLX-100098854) and President Scholarships from Aberystwyth University. The authors would like to thank S. Rahmani for sharing the modified IHS method's code and anonymous reviewers for their constructive comments that have improved the quality and readability of the paper.

References

[1] A. Garzelli, F. Nencini, Interband structure modeling for pan-sharpening of very high-resolution multispectral images, Information Fusion 6 (3) (2005) 213–224.

[2] J. Liu, J. Huang, S. Liu, H. Li, Q. Zhou, J. Liu, Human visual system consistent quality assessment for remote sensing image fusion, ISPRS Journal of Photogrammetry and Remote Sensing 105 (2015) 79–90.

[3] M. Ghahremani, H. Ghassemian, Remote-sensing image fusion based on curvelets and ica, International Journal of Remote Sensing 36 (16) (2015) 4131–4143.

[4] S. Li, X. Kang, L. Fang, J. Hu, H. Yin, Pixel-level image fusion: A survey of the state of the art, Information Fusion 33 (2017) 100–112.

[5] H. Yesou, Y. Besnus, J. Rolet, Extraction of spectral information from landsat tm data and merger with spot panchromatic imageryâĂŤa contribution to the study of geological structures, ISPRS Journal of Photogrammetry and Remote Sensing 48 (5) (1993) 23–36.

[6] L. Eikvil, L. Aurdal, H. Koren, Classification-based vehicle detection in high-resolution satellite images, ISPRS Journal of Photogrammetry and Remote Sensing 64 (1) (2009) 65–72.

[7] T. R. Martha, N. Kerle, V. Jetten, C. J. van Westen, K. V. Kumar, Characterising spectral, spatial and morphometric properties of landslides for semi-automatic detection using object-oriented methods, Geomorphology 116 (1-2) (2010) 24–36.

[8] T. R. Martha, N. Kerle, C. J. van Westen, V. Jetten, K. V. Kumar, Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories, ISPRS Journal of Photogrammetry and Remote Sensing 67 (2012) 105–119.

[9] I. Grinias, C. Panagiotakis, G. Tziritas, Mrf-based segmentation and unsupervised classification for building and road detection in peri-urban areas of high-resolution satellite images, ISPRS Journal of Photogrammetry and Remote Sensing 122 (2016) 145–166.

[10] S. Hasan, C. Montzka, C. Rüdiger, M. Ali, H. R. Bogena, H. Vereecken, Soil moisture retrieval from airborne l-band passive microwave using high resolution multispectral data, ISPRS Journal of Photogrammetry and Remote Sensing 91 (2014) 59–71.

[11] T.-M. Tu, S.-C. Su, H.-C. Shyu, P. S. Huang, A new look at ihs-like image fusion methods, Information fusion 2 (3) (2001) 177–186.

[12] C. A. Laben, B. V. Brower, Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening, uS Patent 6,011,875 (Jan. 4 2000).

[13] P. Chavez, S. C. Sides, J. A. Anderson, et al., Comparison of three different methods to merge multiresolution and multispectral data- landsat tm and spot panchromatic, Photogrammetric Engineering and Remote Sensing 57 (3) (1991) 295–303.

[14] S. Li, J. T. Kwok, Y. Wang, Using the discrete wavelet frame transform to merge landsat tm and spot panchromatic images, Information Fusion 3 (1) (2002) 17–23.

[15] J. J. Lewis, R. J. OâĂŹCallaghan, S. G. Nikolov, D. R. Bull, N. Canagarajah, Pixel-and region-based image fusion with complex wavelets, Information Fusion 8 (2) (2007) 119–130.

[16] M. González-Audícana, X. Otazu, O. Fors, A. Seco, Comparison between mallat's and the âĂŸà trousâĂŹ discrete wavelet transform based algorithms for the fusion of multispectral and panchromatic images, International Journal of Remote Sensing 26 (3) (2005) 595–614.

[17] F. Nencini, A. Garzelli, S. Baronti, L. Alparone, Remote sensing image fusion using the curvelet transform, Information Fusion 8 (2) (2007) 143–156.

[18] S. Yang, M. Wang, L. Jiao, Fusion of multispectral and panchromatic images based on support value transform and adaptive principal component analysis, Information Fusion 13 (3) (2012) 177–184.

[19] S. Yang, M. Wang, L. Jiao, R. Wu, Z. Wang, Image fusion based on a new contourlet packet, Information Fusion 11 (2) (2010) 78–84.

[20] D. Anandhi, S. Valli, An algorithm for multi-sensor image fusion using maximum a posteriori and nonsubsampled contourlet transform, Computers & Electrical Engineering 65 (2018) 139 – 152.

[21] M. Ghahremani, H. Ghassemian, Remote sensing image fusion using ripplet transform and compressed sensing, IEEE Geoscience and Remote Sensing Letters 12 (3) (2015) 502–506.

[22] B. Aiazzi, S. Baronti, F. Lotti, M. Selva, A comparison between global and context-adaptive pansharpening of multispectral images, IEEE Geoscience and Remote Sensing Letters 6 (2) (2009) 302–306.

[23] M. Ghahremani, H. Ghassemian, Nonlinear ihs: A promising method for pan-sharpening, IEEE Geoscience and Remote Sensing Letters 13 (11) (2016) 1606–1610.

[24] J. Cheng, H. Liu, T. Liu, F. Wang, H. Li, Remote sensing image fusion via wavelet transform and sparse representation, ISPRS Journal of Photogrammetry and Remote Sensing 104 (2015) 158–173.

[25] N. Audebert, B. L. Saux, S. LefÃÍvre, Beyond rgb: Very high resolution urban remote sensing with multimodal deep networks, ISPRS Journal of Photogrammetry and Remote Sensing 140 (2018) 20 – 32.

[26] S. Wang, D. Quan, X. Liang, M. Ning, Y. Guo, L. Jiao, A deep learning framework for remote sensing image registration, ISPRS Journal of Photogrammetry and Remote Sensing (2018) In Press.

[27] Y. Xing, M. Wang, S. Yang, L. Jiao, Pan-sharpening via deep metric learning, ISPRS Journal of Photogrammetry and Remote Sensing (2018) In Press.

[28] J. Duran, A. Buades, B. Coll, C. Sbert, G. Blanchet, A survey of pansharpening methods with a new band-decoupled variational model, ISPRS Journal of Photogrammetry and Remote Sensing 125 (2017) 78–105.

[29] H. Ghassemian, A review of remote sensing image fusion methods, Information Fusion 32 (2016) 75–89.

[30] Q. Zhang, Y. Liu, R. S. Blum, J. Han, D. Tao, Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: A review, Information Fusion 40 (2018) 57–75.

[31] Z. Li, H. Leung, Fusion of multispectral and panchromatic images using a restoration-based method, IEEE Transactions on Geoscience and Remote Sensing. 47 (5) (2009) 1482–1491.

[32] F. Palsson, J. R. Sveinsson, M. O. Ulfarsson, A new pansharpening algorithm based on total variation, IEEE Geoscience and Remote Sensing Letters 11 (1) (2014) 318–322.

[33] H. A. Aly, G. Sharma, A regularized model-based optimization framework for pan-sharpening, IEEE Transactions on Image Processing 23 (6) (2014) 2596–2608.

[34] E. J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, IEEE Transactions on Information Theory 52 (2) (2006) 489–509.

[35] D. L. Donoho, Compressed sensing, IEEE Transactions on Information Theory 52 (4) (2006) 1289–1306.

[36] J. Hahn, C. Debes, M. Leigsnering, A. M. Zoubir, Compressive sensing and adaptive direct sampling in hyperspectral imaging, Digital Signal Processing 26 (2014) 113–126.

[37] L. Zhang, W. Wei, Y. Zhang, C. Tian, F. Li, Reweighted laplace prior based hyperspectral compressive sensing for unknown sparsity, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 2274–2281.

[38] Y. Fu, Y. Zheng, I. Sato, Y. Sato, Exploiting spectral-spatial correlation for coded hyperspectral image restoration, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 3727–3736.

[39] J. M. Nascimento, G. Martin, Hyperspectral compressive sensing on low energy consumption board, in: IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2018, pp. 5065–5068.

[40] H. Liu, R. Chen, Y. Wang, P. Lv, Fast omp reconstruction for compressive hyperspectral imaging using joint spatial-spectral sparsity model, in: Tenth International Conference on Information Optics and Photonics, Vol. 10964, International Society for Optics and Photonics, 2018.

[41] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, W.-K. Ma, Hyperspectral super-resolution: Combining low rank tensor and matrix structure, in: 2018 25th IEEE International Conference on Image Processing (ICIP), IEEE, 2018, pp. 3318–3322.

[42] S. Li, B. Yang, A new pan-sharpening method using a compressed sensing technique, IEEE Transactions on Geoscience and Remote Sensing 49 (2) (2011) 738–746.

[43] C. Jiang, H. Zhang, H. Shen, L. Zhang, A practical compressed sensing-based pan-sharpening method, IEEE Geoscience and Remote Sensing Letters 9 (4) (2012) 629–633.

[44] S. Li, H. Yin, L. Fang, Remote sensing image fusion via sparse representations over learned dictionaries, IEEE Transactions on Geoscience and Remote Sensing 51 (9) (2013) 4779–4789.

[45] M. Cheng, C. Wang, J. Li, Sparse representation based pansharpening using trained dictionary, IEEE Geoscience and Remote Sensing Letters 11 (1) (2014) 293–297.

[46] W. Wang, L. Jiao, S. Yang, Fusion of multispectral and panchromatic images via sparse representation and local autoregressive model, Information Fusion 20 (2014) 73–87.

[47] X. X. Zhu, R. Bamler, A sparse image fusion algorithm with application to pan-sharpening, IEEE Transactions on Geoscience and Remote Sensing 51 (5) (2013) 2827–2836.

[48] C. Jiang, H. Zhang, H. Shen, L. Zhang, Two-step sparse coding for the pan-sharpening of remote sensing images, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7 (5) (2014) 1792–1805.

[49] W. Wang, L. Jiao, S. Yang, K. Rong, Distributed compressed sensing-based pan-sharpening with hybrid dictionary, Neurocomputing 155 (2015) 320–333.

[50] H. Yin, Sparse representation based pansharpening with details injection model, Signal Processing 113 (2015) 218–227.

[51] M. Aharon, M. Elad, A. Bruckstein, *rmk*-svd: An algorithm for designing overcomplete dictionaries for sparse representation, IEEE Transactions on Signal Processing 54 (11) (2006) 4311–4322.

[52] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, et al., Distributed optimization and statistical learning via the alternating direction method of multipliers, Foundations and Trends® in Machine Learning 3 (1) (2011) 1–122.

[53] L. Wald, T. Ranchin, M. Mangolini, Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images, Photogrammetric Engineering and Remote Sensing 63 (6) (1997) 691–699.

[54] D. Glasner, S. Bagon, M. Irani, Super-resolution from a single image, in: Computer Vision, 2009 IEEE 12th International Conference on, IEEE, 2009, pp. 349–356.

[55] R. Tibshirani, Regression shrinkage and selection via the lasso, Journal of the Royal Statistical Society. Series B (Methodological) (1996) 267–288.

[56] L. Alparone, S. Baronti, A. Garzelli, F. Nencini, A global quality measurement of pan-sharpened multispectral imagery, IEEE Geoscience and Remote Sensing Letters 1 (4) (2004) 313–317.

[57] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, M. Selva, Multispectral and panchromatic data fusion assessment without reference, Photogrammetric Engineering & Remote Sensing 74 (2) (2008) 193–200.

[58] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, T. Wittman, An adaptive ihs pan-sharpening method, IEEE Geoscience and Remote Sensing Letters 7 (4) (2010) 746–750.

[59] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J.-Y. Tourneret, S. Godsill, Blind model-based fusion of multi-band and panchromatic images, in: Multisensor Fusion and Integration for Intelligent Systems (MFI), 2016 IEEE International Conference on, IEEE, 2016, pp. 21–25.

[60] Q. Wei, N. Dobigeon, J.-Y. Tourneret, J. Bioucas-Dias, S. Godsill, R-fuse: Robust fast fusion of multiband images based on solving a sylvester equation, IEEE Signal Processing Letters 23 (11) (2016) 1632–1636.

[61] R. Restaino, M. Dalla Mura, G. Vivone, J. Chanussot, Context-adaptive pansharpening based on image segmentation, IEEE Transactions on Geoscience and Remote Sensing 55 (2) (2017) 753–766.

Cranfield University

CERES Research Repository **https://dspace.lib.cranfield.ac.uk/**

Cranfield Defence and Security Staff publications (CDS)

Remote sensing image fusion via compressive sensing

Ghahremani, Morteza

2019-04-05

Attribution-NonCommercial 4.0 International

Ghahremani M, Liu Y, Yuen P, Behera A. Remote sensing image fusion via compressive sensing. ISPRS Journal of Photogrammetry and Remote Sensing, Volume 152, June 2019, pp. 34-48
https://doi.org/10.1016/j.isprsjprs.2019.04.001

Downloaded from CERES Research Repository, Cranfield University