Frontiers of Information Technology & Electronic Engineering
www.zju.edu.cn/jzus; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Gradient-based compressive image fusion*

Yang CHEN‡, Zheng QIN
(*Department of Computer Science & Technology, Tsinghua University, Beijing 100084, China*)
E-mail: yang-chen07@mails.tsinghua.edu.cn; qingzh@mail.tsinghua.edu.cn

Received June 18, 2014; Revision accepted Oct. 30, 2014; Crosschecked Jan. 28, 2015

Abstract: We present a novel image fusion scheme based on gradient and scrambled block Hadamard ensemble (SBHE) sampling for compressive sensing imaging. First, source images are compressed by compressive sensing, to facilitate the transmission of the sensor. In the fusion phase, the image gradient is calculated to reflect the abundance of its contour information. By compositing the gradient of each image, gradient-based weights are obtained, with which compressive sensing coefficients are achieved. Finally, inverse transformation is applied to the coefficients derived from fusion, and the fused image is obtained. Information entropy (IE), Xydeas's and Piella's metrics are applied as non-reference objective metrics to evaluate the fusion quality in line with different fusion schemes. In addition, different image fusion application scenarios are applied to explore the scenario adaptability of the proposed scheme. Simulation results demonstrate that the gradient-based scheme has the best performance, in terms of both subjective judgment and objective metrics. Furthermore, the gradient-based fusion scheme proposed in this paper can be applied in different fusion scenarios.

Key words: Compressive sensing (CS), Image fusion, Gradient-based image fusion, CS-based image fusion
doi:10.1631/FITEE.1400217 **Document code:** A **CLC number:** TP391

1 Introduction

Image fusion is a process of combining information from multiple images into a single fused image. According to the development of fusion, image fusion technologies can be classified into three hierarchical levels, pixel-level (Yang Y *et al.*, 2007), feature-level (Ross and Govindarajan, 2005), and decision-level (Byeungwoo and Landgrebe, 1999). Pixel-level fusion refers to image processing directly based on the original pixel's information of the source images. Since real-world objects usually consist of structures at different scales, some researchers have studied various multiresolution (multiscale) representations of a signal at the pixel level. A multiresolution decomposition scheme decomposes the signal being analyzed into several components, each of which captures information representation at a given scale. Therefore, multiresolution information is useful in image fusion. These various multiresolution transforms include pyramid (Liu *et al.*, 2001), gradient (Petrović and Xydeas, 2004), wavelet (Pajares and de la Cruz, 2004; Shi *et al.*, 2005; Amolins *et al.*, 2007), and bidimensional empirical mode (Zheng and Qin, 2009).

In recent years, with the rising attention to compressive sensing (CS) theory (Donoho, 2006; Candès and Wakin, 2008) from academic and industrial worlds, image fusion for compressive sensing has attracted the attention of many researchers. Compared with the traditional image fusion which requires the whole acquisition of the source images, compressive sensing image fusion does not have any requirement on the source image samplings. CS ensures that if a signal is sparse on a certain basis, it can be recovered from a relatively small set of random linear projections on another basis which is incoherent

‡ Corresponding author
* Project supported by the National S&T Major Program (No. 9140A1550212 JW01047) and the 'Twelfth Five' Preliminary Research Project of PLA (No. 402040202)
ORCID: Yang CHEN, http://orcid.org/0000-0003-0927-000X

with the sparse basis. Therefore, with the superiority in reducing computational and transmission costs, CS has become a much preferred algorithm for image fusion.

Previous literature on CS image fusion explores the fusion schemes based on the sampling model. Different fusion schemes have been conducted based on different sampling models. Wan and Qin (2011) explored the compressive image fusion by using the Fourier transform as its sparse basis, and the maximum of absolute values (MAV) as the fusion scheme. Luo *et al.* (2009) took the scrambled block Hadamard ensemble (SBHE) as the sampling operator, and proposed a fusion algorithm in which the weights are calculated based on entropy metrics of original measurements. Han *et al.* (2010) used discrete cosine transform (DCT) as the sparse basis. Measurements are conducted with a wavelet transform, and then the approximation coefficients and detail coefficients are fused separately. Kang *et al.* (2013) used the dual-channel pulse coupled neural network model in the image sampling part as an important weight factor in the fusion scheme. Ding *et al.* (2013) selected three trained over-complete dictionaries by *K*-means singular value decomposition (*K*-SVD) including the dictionary using only patches from the infrared images, the dictionary using only patches from visible images, and the dictionary using the combined patches, and two fusion rules covering the maximum l_1-norm and the maximum absolute for entry of the sparse vector.

There are different ensembles of CS matrices defined in the literature (Duarte *et al.*, 2008; Romberg, 2008; Do *et al*., 2012). The SBHE sampling operator (Do *et al.*, 2012) employed in this paper has very good properties and it can be easily applied to the optical domain (such as single-pixel cameras), requiring small storage space but with high computing speed. In addition, in contrast with the presented CS fusion literature, this study proposes a novel gradient-based image fusion framework, which considers the abundance of its contour information.

Since multiple image sensors are widely employed in many fields such as multi-focus, military, and medical imaging, to increase the capabilities of intelligent machines and systems, it is necessary to explore the scenario adaptabilities of different fusion schemes for different scenarios. For example, imaging cameras usually have only a finite depth of field. Only those objects within the depth of field of the camera are focused, while other objects are blurred. Therefore, multi-focus image fusion is used to create a better description of the scene compared with any of the individual focused images. In the infrared (IR) and visible image fusion scenario, an IR image is sensitive to IR light with low definition, while a visible image is captured with more details of the scene. Thus, the fusion result of IR and visible images can obtain a comprehensive representation of both important objects detected by an IR image and the environmental details from visible images.

2 Related studies

2.1 Compressive sensing theory

Compressive sensing theory (Donoho, 2006; Candès and Wakin, 2008) believes that using less sampling or measurement compared with the traditional method can restore signals or images. The unique part of CS sampling is that it allows the sensor to capture the signal from sparse signals efficiently, without trying to understand the signal. With just a small amount of the acquired data, all signals can be reconstructed using the numerical optimization method at the receiving end.

The core idea of the CS theory includes mainly two parts. The first is the sparse signal structure. The traditional representation method of the Shannon signal develops and utilizes only the minimum amount of prior information of sampled signals, namely the bandwidth of the signal. However, in real life, we pay much more attention to popular signals which have some structural characteristics themselves. Corresponding to the freedom degree of bandwidth information, the structural characteristics are determined by a smaller fraction of signal freedom degrees. In other words, in the rare information loss circumstance, signal can be represented by fewer digital codes. Therefore, in this sense, the signal is a sparse signal (or approximately sparse signals, compressible signals). Another point is the characteristic of no correlation. Useful information for sparse signals can be obtained with a sampling combined with a compression method for non-adaptive signals, which compresses the signal into a small sample data.

Theories proved that the sampling method of compressed sensing is only a simple operation correlating the signal with a set of established waves. The waveform is not necessarily associated with the signal in sparse space.

Given signal $\boldsymbol{x}\in\mathbb{R}^N$, it is expressed in the following form on the orthogonal base:

$$\boldsymbol{x}=\sum_{i=0}^{N-1}a_i\boldsymbol{\psi}_i=\boldsymbol{\psi a}. \tag{1}$$

For the coefficient vector $\boldsymbol{a}$, if only K elements are not zero, we say K ($K\geq1$) is sparse when $\boldsymbol{x}$ is in the substrate $\boldsymbol{\Psi}$. After the measurement matrix $\boldsymbol{M}\in\mathbb{R}^{M\times N}$ projection, there is

$$\boldsymbol{y}=\boldsymbol{Mx}=\boldsymbol{M\Psi a}. \tag{2}$$

Because $M\ll N$, the projection process combines the traditional signal sampling and compression processes. Through $\boldsymbol{y}$ and $\boldsymbol{M}$, direct recovery of $\boldsymbol{x}$ is a morbid inverse problem. However, sparse theory suggests that this problem can be transformed into an optimization problem of sparse vectors to be solved:

$$\min\|\boldsymbol{a}\|_1 \text{ s.t. } \boldsymbol{y}=\boldsymbol{\Phi a}, \tag{3}$$

where $\boldsymbol{\Phi}=\boldsymbol{M\Psi}$, $\|\cdot\|_1$ represents the l_1-norm. Thus, signal $\boldsymbol{x}$ can be restored from formula $\boldsymbol{x}=\boldsymbol{\Psi a}$.

Subsequently, Candès and Wakin (2008) pointed out that if we want to accurately recover signal $\boldsymbol{x}$, the matrix $\boldsymbol{\Phi}$ must satisfy the restricted isometry property (RIP):

$$1-\varepsilon\leq\frac{\|\boldsymbol{\Phi v}\|_2}{\|\boldsymbol{v}\|_2}\leq1+\varepsilon, \tag{4}$$

where $\varepsilon>0$, $\|\cdot\|_2$ represents the 2-norm, and $\boldsymbol{v}$ is an N-dimensional vector with strict K sparseness. That is to say, matrix $\boldsymbol{M}$ and substrate $\boldsymbol{\Psi}$ are not related.

2.2 Scrambled block Hadamard ensemble sampling

Divide an image into small blocks, where the block size is $B\times B=N$. Corresponding to the measurement matrix of the ith block is $\boldsymbol{\Phi}_i$, whose size is $m_i\times N$. All measurement matrices are generated by the SBHE operator (Do *et al.*, 2012):

$$\boldsymbol{\Phi}_i=\boldsymbol{Q}_{m_i}\boldsymbol{HP}_N, \tag{5}$$

where $\boldsymbol{P}_N$ represents an N-column vector derived from the random permutation of $\boldsymbol{H}$, $\boldsymbol{Q}_{m_i}$ represents an m_i-row vector derived from the random permutation of $\boldsymbol{HP}_N$, whereas $\boldsymbol{H}$ represents the block diagonal matrix:

$$\boldsymbol{H}=\begin{bmatrix}\boldsymbol{H}_1 & & & \\ & \boldsymbol{H}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{H}_B\end{bmatrix}. \tag{6}$$

$\boldsymbol{H}_j\in\mathbb{R}^{B\times B}$ ($1\leq j\leq B$) represents the Hadamard matrices. The SBHE operator has many excellent properties: First, it is a universal measurement matrix, which can accurately reconstruct the original signal. Second, it can quickly generate Hadamard matrices. Finally, the randomness of $\boldsymbol{P}_N$ and $\boldsymbol{Q}_{m_i}$ can reduce the blocking effect of the reconstructed signal.

Then vectorize the ith image block by column scanning and obtain $\boldsymbol{x}_i$, whose corresponding measurement output vector is expressed as

$$\boldsymbol{y}_i=\boldsymbol{\Phi}_i\boldsymbol{x}_i. \tag{7}$$

The measurement output vector of the entire image is determined by

$$\boldsymbol{y}=\begin{bmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\\ \vdots\\ \boldsymbol{y}_I\end{bmatrix}. \tag{8}$$

The fused image can be reconstructed from the measurement according to a recovery algorithm such as the gradient projection for sparse reconstruction (GPSR) (Figueiredo *et al.*, 2007), basis pursuit (Chen *et al.*, 1998), total variation minimization (Wang *et al.*, 2008), orthogonal matching pursuit (OMP) (Tropp and Gilbert, 2007), and l_1-norm minimization (Candès and Romberg, 2005; Candès and Tao, 2006).

In this study, the original image is reconstructed using the GPSR algorithm (Figueiredo *et al.*, 2007).

3 Compressive image fusion schemes

3.1 Compressive image fusion framework

The purpose of image fusion is to composite scenes, which contributes to a better understanding of the scene. Generally speaking, a satisfying image fusion approach is required to meet the following two requirements: First, image fusion processing, especially in the sensor side, requires a simple and effective approach, in both image transmission and computation complexities. Second, it is required to be robust, reliable, and tolerant of imperfections such as noise. In this way, an image fusion framework for compressive imaging is proposed (Fig. 1).

First, the source image is compressed through compressive sensing so as to facilitate the transmission of the sensor. In the fusion phase, fusion rules are adopted, making the fusion coefficients much flatter in the region. Then inverse transformation is applied to the coefficients derived from the fusion, and the fused image is eventually obtained.

Most literature uses multi-resolution decompositions and weighted combination image fusion schemes, like average-based fusion, principal component analysis (PCA) based fusion, mutual information (MI) based fusion, and pulse coupled neural network (PCNN) based fusion. Average-based fusion is a simple operational method, which assumes that source images have the same weight in the fusion image. PCA-based fusion can be used to optimize the weights with respect to information content and redundancy removal. The MI-based fusion scheme calculates the weights based on the entropy metrics of the source images. The neural network fusion scheme, such as PCNN, benefits image fusion by using local image information.

In the compressive sensing image fusion framework, we implement the fusion schemes based on SBHE sampling and adopt weighted combination image fusion schemes like average-, mean-, variance-, PCA-, and MI-based fusion, which are used at the picture (pixel) level. These schemes directly combine compressive sensing coefficients, and details are described in the following sections. In addition, gradient-based compressive image fusion by using SBHE is proposed and corresponding experiments are conducted.

3.2 Fusion method based on average

Suppose $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are the measurement output vectors of the jth block of two images, respectively. The mixed measurement output after fusion can be obtained through the weighted mean:

$$\boldsymbol{y}_j=w_1\boldsymbol{v}_1+w_2\boldsymbol{v}_2, \tag{9}$$

where w_1 and w_2 represent the weights. By applying average-based fusion, it is known that the jth blocks of the two images have the same weight, i.e., $w_1=w_2=0.5$.

3.3 Fusion method based on mean

In the fusion method based on mean, the weights are calculated according to

$$w_1=\frac{m_1}{m_1+m_2}, \tag{10}$$

$$w_2=\frac{m_2}{m_1+m_2}, \tag{11}$$

where m_i (i=1, 2) is defined as the first moment of vector $\boldsymbol{v}_i$:

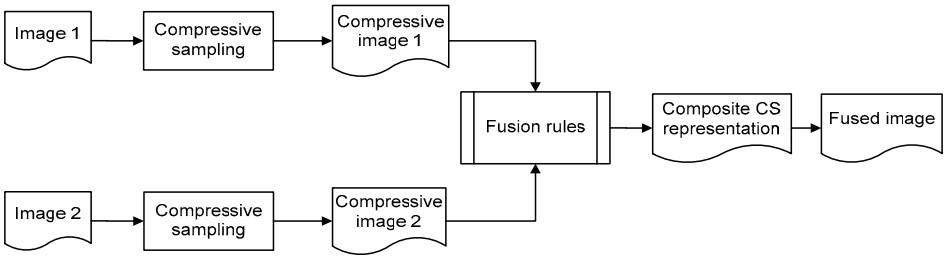

Fig. 1 Compressive image fusion framework

$$m_i=\frac{\sum_k \boldsymbol{v}_i(k)}{\mathrm{card}(\boldsymbol{v}_i)},\quad i\in\{1,2\}. \tag{12}$$

$\boldsymbol{v}_i(k)$ represents the kth element of vector $\boldsymbol{v}_i$ and card(·) is used to calculate the number of elements in the vector.

3.4 Fusion method based on variance

In the fusion method based on variance, the weights are calculated according to

$$w_1=\frac{s_1}{s_1+s_2}, \tag{13}$$

$$w_2=\frac{s_2}{s_1+s_2}, \tag{14}$$

where s_i (i=1, 2) is defined as

$$s_i=\frac{\sum_k[\boldsymbol{v}_i(k)-m_i]^2}{\mathrm{card}(\boldsymbol{v}_i)-1},\quad i\in\{1,2\}. \tag{15}$$

3.5 Fusion method based on information theory

Mutual information (Qu *et al*., 2002) is a useful measure in information theory which refers to the correlation between two event sets. Mutual information has been widely used in image processing.

The information entropy, joint information entropy, and mutual information of the two measurement outputs should be calculated, in order to compute weight w_i:

$$H(\boldsymbol{v}_i)=-\sum_{v_i}P(\boldsymbol{v}_i)\ln P(\boldsymbol{v}_i),\ i\in\{1,2\}, \tag{16}$$

$$H(\boldsymbol{v}_1,\boldsymbol{v}_2)=-\sum_{v_1}\sum_{v_2}P(\boldsymbol{v}_1,\boldsymbol{v}_2)\ln P(\boldsymbol{v}_1,\boldsymbol{v}_2), \tag{17}$$

$$I(\boldsymbol{v}_1,\boldsymbol{v}_2)=\sum_{v_1}\sum_{v_2}P(\boldsymbol{v}_1,\boldsymbol{v}_2)\ln\frac{P(\boldsymbol{v}_1,\boldsymbol{v}_2)}{P(\boldsymbol{v}_1)P(\boldsymbol{v}_2)}. \tag{18}$$

The weights can be calculated according to

$$w_1=\frac{H(\boldsymbol{v}_1)}{H(\boldsymbol{v}_1,\boldsymbol{v}_2)}-\frac{I(\boldsymbol{v}_1,\boldsymbol{v}_2)}{2H(\boldsymbol{v}_1)}\cdot\frac{I(\boldsymbol{v}_1,\boldsymbol{v}_2)}{H(\boldsymbol{v}_1,\boldsymbol{v}_2)}, \tag{19}$$

$$w_2=\frac{H(\boldsymbol{v}_2)}{H(\boldsymbol{v}_1,\boldsymbol{v}_2)}-\frac{I(\boldsymbol{v}_1,\boldsymbol{v}_2)}{2H(\boldsymbol{v}_2)}\cdot\frac{I(\boldsymbol{v}_1,\boldsymbol{v}_2)}{H(\boldsymbol{v}_1,\boldsymbol{v}_2)}. \tag{20}$$

The first term on the right of Eq. (19) describes the weight of information entropy in the joint information entropy. The sole use of the first term will lead to the exposure of the image after fusion. This is because the first term has calculated the common part of the two images twice. The revised w_1 can be obtained if we deduct the second term, which is the common part of the two images. Finally, we will integrate w_1 with w_2, i.e., making $w_1+w_2=1$.

If $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are independent, $I(\boldsymbol{v}_1,\boldsymbol{v}_2)=0$. Then the weight is the proportion of the respective information entropy in the joint information entropy. If $\boldsymbol{v}_1=\boldsymbol{v}_2$, then $w_1=w_2$.

3.6 Fusion method based on PCA

PCA (Jolliffe, 1986; Smith, 2002) is a statistical analysis method used to grasp the principal contradiction of things. It can be used to analyze the main factors of diverse things, revealing the essence of things and simplifying complex issues. Due to the large dimension of image information, PCA has been widely used in image processing for data dimension reduction.

The PCA-based weight calculation method uses the two sub-blocks' measurement outputs $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ to make a sample matrix:

$$\boldsymbol{V}=\begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2\end{bmatrix}. \tag{21}$$

The covariance matrix of $\boldsymbol{V}$ is calculated as follows:

$$\boldsymbol{C}=\frac{1}{M-1}\sum_{i=0}^{M}(\boldsymbol{V}_i-\boldsymbol{m})^{\mathrm{T}}(\boldsymbol{V}_i-\boldsymbol{m}), \tag{22}$$

$$\boldsymbol{m}=\frac{1}{M}\sum_{i=0}^{M}\boldsymbol{V}_i, \tag{23}$$

where M is the total number of samples (each row represents one sample), $\boldsymbol{V}_i$ represents the ith row sample of $\boldsymbol{V}$, and $\boldsymbol{m}$ represents the sample mean. The eigenvalue of $\boldsymbol{C}$ can be calculated by

$$|\lambda\boldsymbol{I}-\boldsymbol{C}|=0, \tag{24}$$

where $\boldsymbol{I}$ represents the identity matrix. The maximum eigenvalue $\lambda_{\max}$ and its corresponding eigenvector $\boldsymbol{u}$ satisfy the following relation:

$$\boldsymbol{C}\boldsymbol{u}=\lambda\boldsymbol{u}. \tag{25}$$

The corresponding eigenvector $\boldsymbol{u}$ is actually the optimal projection direction of the sample space. Thus,

the weights for measurement of outputs $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are defined as

$$w_1=\frac{u_1}{u_1+u_2}, \tag{26}$$

$$w_2=\frac{u_2}{u_1+u_2}, \tag{27}$$

where u_i (i=1, 2) represents the ith component of the eigenvector.

The PCA-based image fusion method can remove the redundant information of the source image, and gain the principal component weight and corresponding eigenvalue of each source image. The weight is determined by the corresponding eigenvalue of the source image.

3.7 Fusion method based on gradient

Let $I_1(x, y)$ and $I_2(x, y)$ represent the jth block in image 1 and image 2, respectively. Their Gaussian gradients are defined as follows:

$$\nabla I_i(x, y)=\{I_i\cdot\nabla G_\sigma\}(x, y), \tag{28}$$

$$G_\sigma(x, y)=\frac{1}{2\pi\sigma}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \tag{29}$$

where $i\in\{1, 2\}$, $G_\sigma(x, y)$ is the Gaussian function for smoothing the image to reduce the effects of noise, and σ is the standard deviation of noise, controlling smoothness. Gradient magnitudes m_1 and m_2 are calculated as follows:

$$m_1=\sum_{x, y}\|\nabla I_1(x, y)\|_1, \tag{30}$$

$$m_2=\sum_{x, y}\|\nabla I_2(x, y)\|_1. \tag{31}$$

Gradient magnitude reflects the abundance of contour information in the image. Blend weights of measurement outputs $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are defined as follows:

$$w_1=\frac{m_1}{m_1+m_2}, \tag{32}$$

$$w_2=\frac{m_2}{m_1+m_2}. \tag{33}$$

The gradient-based weight definition method can utilize effectively the contour information in the image. Since the image is rich in contour information, its corresponding weight will be amplified in the fusion result that will show extensive contour representation, contributing to accurate image judgment.

4 Experimental results

4.1 Experiment setup

Because image fusion is widely used in military and civil fields, we set up two groups of image fusion for different application scenarios. They are the multi-focus image fusion and the fusion of infrared and visible images, respectively. The block size was set at 32×32. In this way, the block Hadamard sampling matrix can encode pixels with 32×32. The sampling rate was set at 0.5.

Objective assessments were also included in our experiments. In this paper, IE, Xydeas's (Xydeas and Petrović, 2000), and Piella's (Piella and Heijmans, 2003) metrics were used as non-reference objective metrics. Information entropy was generally applied to measure the amount of information. The more information entropy there was, the better the fusion result would be. In addition, Xydeas's and Piella's metrics were applied for assessment of the salient information transferred from the input images to fused images. Piella's metric comprehensively takes the image correlation coefficient, mean luminance, contrast, and edge information into account. The dynamic ranges of three Piella's indexes, Q, Qw, and Qe, are [−1, 1]. The closer the values were to 1, the better the fusion performance was expressed. Codes of this study and experimental results can be obtained on the personal website (http://hi.baidu.com/thufuse).

4.2 Experimental results

4.2.1 Application scenario 1: multi-focus image fusion

Multi-focus image fusion (Li *et al.*, 2004; Chen *et al.*, 2008; Li and Yang, 2008) is used to create a better description of the scene than any of the individual focused images. The obtained fusion image shows all the details of the source image, making it more conducive to the follow-up treatment. For example, with multi-focus image fusion, the speed and accuracy of target recognition can be significantly improved. We implemented CS fusion based on six fusion schemes combined with compressive sensing coefficients, and compared these fusion schemes for three sets of multi-focus image fusion. Experimental

results and objective measurements are given in Tables 1–3 and Figs. 2–4.

We have proposed the compressive image fusion scheme based on gradient and SBHE sampling. By means of IE, Xydeas's, and Piella's metrics, fused images can be measured. Tables 1–3 have demonstrated that in the fusion scenario of multi-focus images, our proposed scheme has the best result from the perspective of both objective assessment and visual perception.

4.2.2 Application scenario 2: infrared and visible light image fusion

In military image fusion of infrared and visible light (Yang XH *et al.*, 2007; Li and Qin, 2011; Wang and Du, 2014), the infrared light has a strong ability in discovering important military targets, while the visible light delivers excellent texture expression. Therefore, image fusion embodies not only the important military targets, but also the good texture expression ability in the fusion results. We implemented

Table 2 Comparison of fusion schemes for multi-focus images (pepsi) among different metrics

Scheme	IE	Xydeas	Q*	Qw*	Qe*
CS_Avg	7.0311	0.1937	0.4406	0.6361	0.3899
CS_Mean	7.0312	0.1939	0.4405	0.6354	0.3876
CS_Var	6.2262	0.1174	0.4028	0.4682	0.2052
CS_PCA	7.0312	0.1932	0.4410	0.6373	0.3918
CS_MI	7.0314	0.1927	0.4415	0.6393	0.3979
Proposed	7.0325	0.1853	0.4562	0.6826	0.5086

* Piella's indexes

Table 1 Comparison of fusion schemes for multi-focus images (clock) among different metrics

Scheme	IE	Xydeas	Q*	Qw*	Qe*
CS_Avg	7.3939	0.4174	0.6133	0.7747	0.4878
CS_Mean	7.3814	0.4200	0.6071	0.7550	0.4648
CS_Var	7.3818	0.4182	0.6071	0.7552	0.4642
CS_PCA	7.3812	0.4183	0.6068	0.7546	0.4642
CS_MI	7.3843	0.4227	0.6085	0.7580	0.4690
Proposed	7.4014	0.5155	0.6230	0.7941	0.5668

* Piella's indexes

Table 3 Comparison of fusion schemes for multi-focus images (peppers) among different metrics

Scheme	IE	Xydeas	Q*	Qw*	Qe*
CS_Avg	7.5424	0.4464	0.6345	0.7823	0.5377
CS_Mean	7.5431	0.4466	0.6347	0.7829	0.5385
CS_Var	7.5433	0.4485	0.6344	0.7825	0.5388
CS_PCA	7.5434	0.4495	0.6349	0.7830	0.5401
CS_MI	7.5430	0.4488	0.6344	0.7825	0.5396
Proposed	7.5488	0.4548	0.6360	0.7857	0.5622

* Piella's indexes

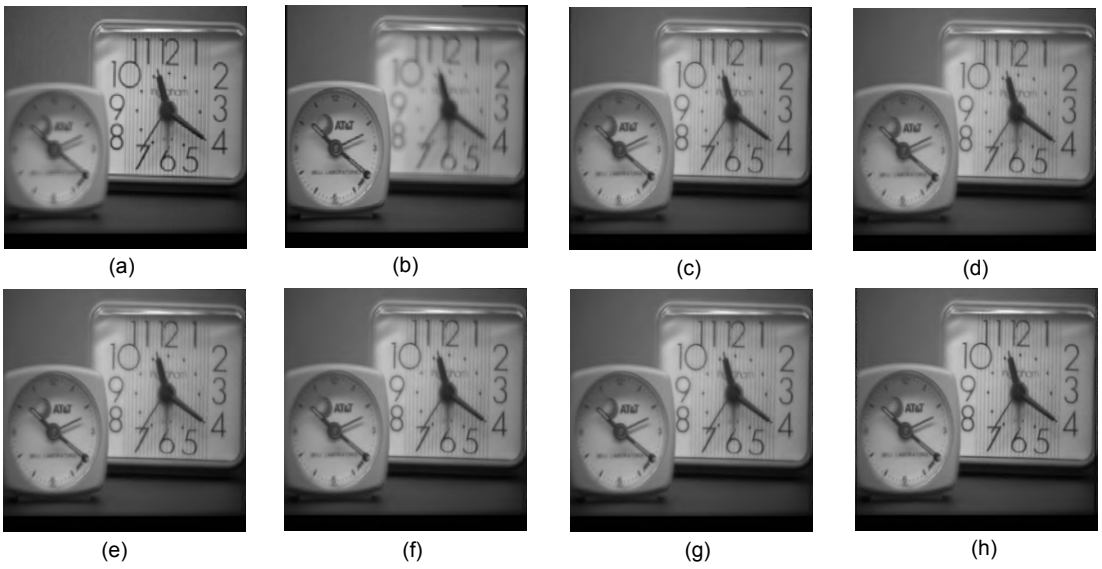

Fig. 2 Multi-focus image fusion (clock)

(a) is the left-focus clock, (b) is the right-focus clock, and (c)–(h) are fusion results under the average scheme, mean scheme, variance scheme, PCA scheme, MI scheme, and the proposed scheme, respectively

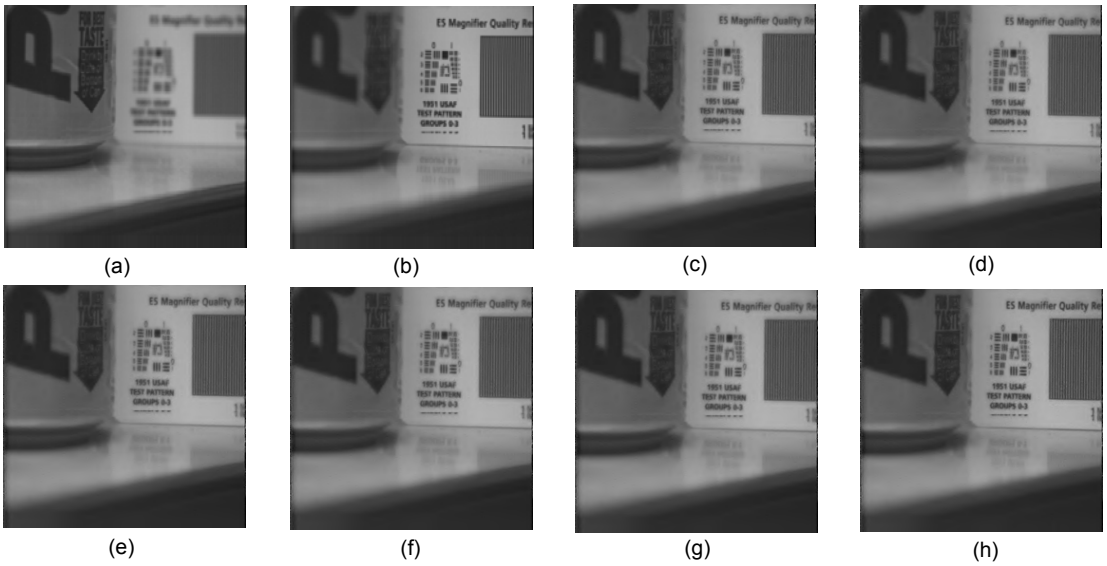

Fig. 3 Multi-focus image fusion (pepsi)

(a) is the left-focus pepsi, (b) is the right-focus pepsi, and (c)–(h) are fusion results under the average scheme, mean scheme, variance scheme, PCA scheme, MI scheme, and the proposed scheme, respectively

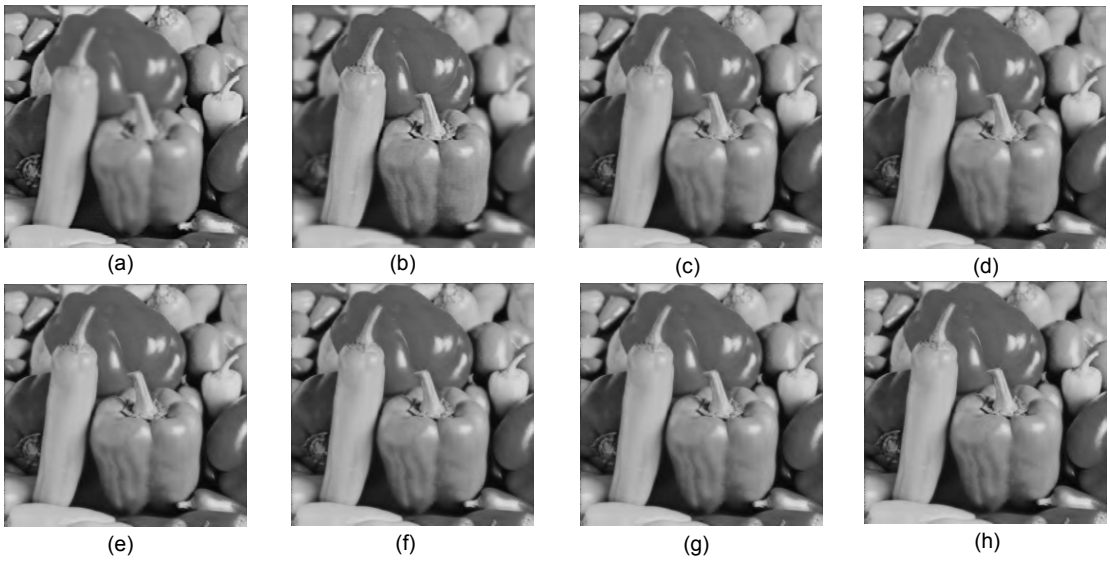

Fig. 4 Multi-focus image fusion (peppers)

(a) is the input image 1 of peppers, (b) is the input image 2 of peppers, and (c)–(h) are fusion results under the average scheme, mean scheme, variance scheme, PCA scheme, MI scheme, and the proposed scheme, respectively

CS image fusion based on six fusion schemes combined with compressive sensing coefficients, and conducted the experiments in UNCAMP image sets (three frames were selected randomly). Experimental results and objective measurements are given in Tables 4–6 and Figs. 5–7.

In infrared and visible light image fusion scenarios (Tables 4–6), the proposed scheme has better performance in IE, Xydeas's, and Piella's metrics than any of the other five fusion schemes. It is also demonstrated that the fused image reserves the edge and salient information from source images.

Based on the feature that human vision is sensitive to local variations, Xydeas and Petrović (2000) and Piella and Heijmans (2003) proposed that fusion image quality can be evaluated by measuring the

similarity between the salient information of the source image and that of the fused image. In this way, tiny changes in edges can cause greater differences in Xydeas's and Piella's metrics. The Xydeas's and Piella's metrics in application scenario 2 could be low due to the changes in the edges between the source images and fused images.

5 Conclusions

This paper proposes a compressive image framework and presents compressive image fusion schemes. It is necessary and valuable to implement possible image fusion schemes and compare them for compressive imaging. From the experimental results, we prove that the gradient-based fusion scheme, proposed in this paper is efficient in terms of both subjective judgement and objective metrics for compressive image fusion. In addition, experiments were conducted under different fusion scenarios. The experimental results demonstrate that this gradient-based fusion scheme is adaptive for different

Table 4 Comparison of fusion schemes for infrared and visible light images (frame 1800) among different metrics

Scheme	IE	Xydeas	Q*	Qw*	Qe*
CS_Avg	6.1771	0.1133	0.4059	0.4581	0.1965
CS_Mean	6.1709	0.1143	0.4038	0.4569	0.1963
CS_Var	6.1828	0.1135	0.4046	0.4573	0.1953
CS_PCA	6.1862	0.1136	0.4049	0.4584	0.1964
CS_MI	6.1752	0.1174	0.4054	0.4578	0.1968
Proposed	6.2033	0.1180	0.4068	0.4625	0.2071

* Piella's indexes

Table 5 Comparison of fusion schemes for infrared and visible light images (frame 1807) among different metrics

Scheme	IE	Xydeas	Q*	Qw*	Qe*
CS_Avg	6.2244	0.1129	0.4059	0.4718	0.2102
CS_Mean	6.2106	0.1165	0.4030	0.4687	0.2066
CS_Var	6.2262	0.1174	0.4028	0.4682	0.2052
CS_PCA	6.2297	0.1187	0.4026	0.4685	0.2060
CS_MI	6.2168	0.1191	0.4021	0.4678	0.2058
Proposed	6.2670	0.1194	0.4067	0.4752	0.2208

* Piella's indexes

Table 6 Comparison of fusion schemes for infrared and visible light images (frame 1831) among different metrics

Scheme	IE	Xydeas	Q*	Qw*	Qe*
CS_Avg	6.2179	0.1085	0.3858	0.4472	0.2060
CS_Mean	6.2183	0.1101	0.3860	0.4482	0.2063
CS_Var	6.2307	0.1118	0.3856	0.4492	0.2085
CS_PCA	6.2354	0.1127	0.3859	0.4501	0.2095
CS_MI	6.2150	0.1149	0.3841	0.4458	0.2059
Proposed	6.2559	0.1170	0.3887	0.4547	0.2218

* Piella's indexes

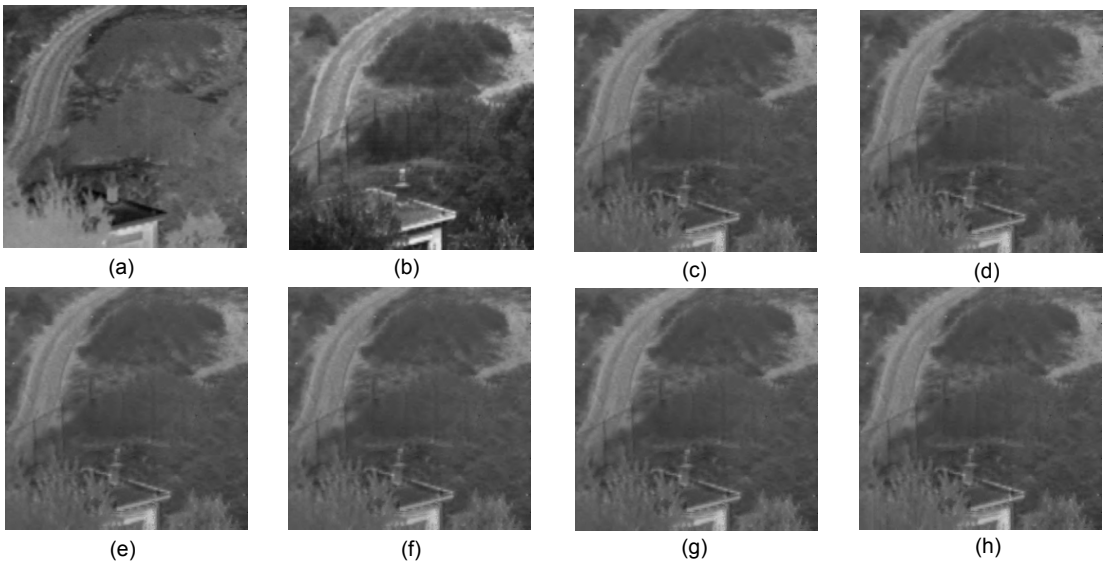

Fig. 5 Image fusion of infrared and visible light (frame 1800)

(a) is the infrared image, (b) is the visible light image, and (c)–(h) are fusion results under the average scheme, mean scheme, variance scheme, PCA scheme, MI scheme, and the proposed scheme, respectively

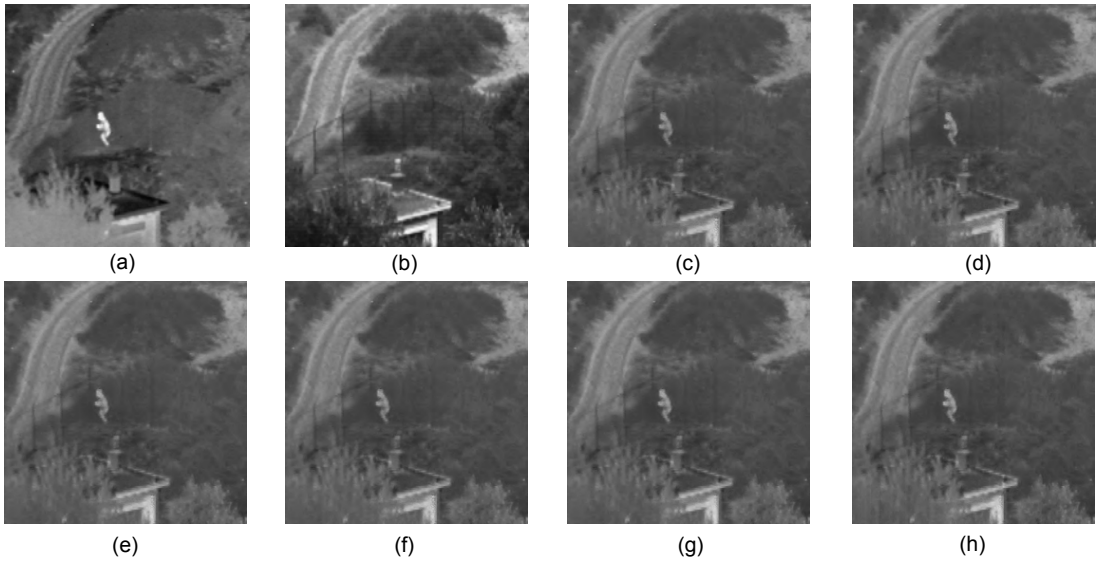

Fig. 6 Image fusion of infrared and visible light (frame 1807)

(a) is the infrared image, (b) is the visible light image, and (c)–(h) are fusion results under the average scheme, mean scheme, variance scheme, PCA scheme, MI scheme, and the proposed scheme, respectively

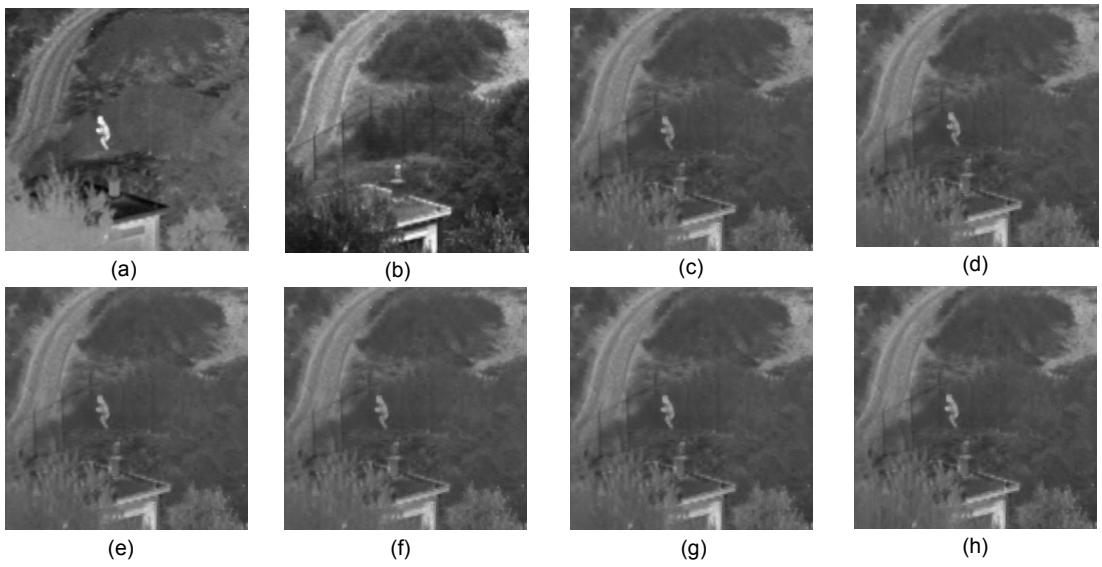

Fig. 7 Image fusion of infrared and visible light (frame 1831)

(a) is the infrared image, (b) is the visible light image, and (c)–(h) are fusion results under the average scheme, mean scheme, variance scheme, PCA scheme, MI scheme, and the proposed scheme, respectively

fusion scenarios. Future work will be focused on exploring gradient-based image fusion schemes by using other compressive sampling and more image fusion schemes will be compared.

References

Amolins, K., Zhang, Y., Dare, P., 2007. Wavelet based image fusion techniques—an introduction, review and comparison. *ISPRS J. Photogram. Remote Sens.*, **62**(4): 249-263. [doi:10.1016/j.isprsjprs.2007.05.009]

Byeungwoo, J., Landgrebe, D.A., 1999. Decision fusion approach for multitemporal classification. *IEEE Trans. Geosci. Remote Sens.*, **37**(3):1227-1233. [doi:10.1109/36.763278]

Candès, E.J., Romberg, J., 2005. l_1-Magic: Recovery of Sparse Signals via Convex Programming. Available from http://www.acm.caltech.edu/l1magic/

Candès, E.J., Tao, T., 2006. Near-optimal signal recovery from random projections: universal encoding strategies. *IEEE Trans. Inform. Theory*, **52**(12):5406-5425. [doi:10.1109/TIT.2006.885507]

Candès, E.J., Wakin, M.B., 2008. An introduction to compressive sampling. *IEEE Signal Process. Mag.*, **25**(2):21-30. [doi:10.1109/MSP.2007.914731]

Chen, R.Y., Li, S., Yang, R., *et al*., 2008. Multi-focus images fusion based on data assimilation and genetic algorithm. Proc. Int. Conf. on Computer Science and Software Engineering, p.249-252. [doi:10.1109/CSSE.2008.525]

Chen, S.S., Donoho, D.L., Saunders, M.A., 1998. Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.*, **20**(1):33-61. [doi:10.1137/S1064827596304010]

Ding, M., Wei, L., Wang, B.F., 2013. Research on fusion method for infrared and visible images via compressive sensing. *Infrared Phys. Technol.*, **57**:56-67. [doi:10.1016/j.infrared.2012.12.014]

Do, T.T., Lu, G., Nguyen, N.H., *et al*., 2012. Fast and efficient compressive sensing using structurally random matrices. *IEEE Trans. Signal Process.*, **60**(1):139-154. [doi:10.1109/TSP.2011.2170977]

Donoho, D.L., 2006. Compressed sensing. *IEEE Trans. Inform. Theory*, **52**(4):1289-1306. [doi:10.1109/TIT.2006.871582]

Duarte, M.F., Davenpot, M.A., Takhar, D., *et al.*, 2008. Single-pixel imaging via compressive sampling. *IEEE Signal Process. Mag*., **25**(2):83-91. [doi:10.1109/MSP.2007.914730]

Figueiredo, M.A.T., Nowak, R.D., Wright, S.J., 2007. Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems. *IEEE J. Sel. Topics Signal Process.*, **1**(4):586-597. [doi:10.1109/JSTSP.2007.910281]

Han, J.J., Loffeld, O., Hartmann, K., *et al.*, 2010. Multi image fusion based on compressive sensing. Proc. Int. Conf. on Audio Language and Image Processing, p.1463-1469. [doi:10.1109/ICALIP.2010.5684502]

Jolliffe, I.T., 1986. Principal Component Analysis. Springer.

Kang, B., Zhu, W.P., Yan, J., 2013. Fusion framework for multi-focus images based on compressed sensing. *IET Image Process.*, **7**(4):290-299. [doi:10.1049/iet-ipr.2012.0543]

Li, S.T., Yang, B., 2008. Multifocus image fusion by combining curvelet and wavelet transform. *Patt. Recog. Lett.*, **29**(9):1295-1301. [doi:10.1016/j.patrec.2008.02.002]

Li, S.T., Kwok, J.T.Y., Tsang, I.W., *et al*., 2004. Fusing images with different focuses using support vector machines. *IEEE Trans. Neur. Netw.*, **15**(6):1555-1561. [doi:10.1109/TNN.2004.837780]

Li, X., Qin, S.Y., 2011. Efficient fusion for infrared and visible images based on compressive sensing principle. *IET Image Process.*, **5**(2):141-147. [doi:10.1049/iet-ipr.2010.0084]

Liu, Z., Tsukada, K., Hanasaki, K., *et al.*, 2001. Image fusion by using steerable pyramid. *Patt. Recog. Lett*., **22**(9):929-939. [doi:10.1016/S0167-8655(01)00047-2]

Luo, X.Y., Zhang, J., Yang, J.Y., *et al.*, 2009. Image fusion in compressed sensing. Proc. 16th IEEE Int. Conf. on Image Processing, p.2205-2208. [doi:10.1109/ICIP.2009.5413866]

Pajares, G., de la Cruz, J.M., 2004. A wavelet-based image fusion tutorial. *Patt. Recog.*, **37**(9):1855-1872. [doi:10.1016/j.patcog.2004.03.010]

Petrović, V.S., Xydeas, C.S., 2004. Gradient-based multiresolution image fusion. *IEEE Trans. Image Process.*, **13**(2):228-237. [doi:10.1109/TIP.2004.823821]

Piella, G., Heijmans, H., 2003. A new quality metric for image fusion. Proc. Int. Conf. on Image Processing, p.173-176. [doi:10.1109/ICIP.2003.1247209]

Qu, G.H., Zhang, D.L., Yan, P.F., 2002. Information measure for performance of image fusion. *Electron. Lett.*, **38**(7):313-315. [doi:10.1049/el:20020212]

Romberg, J., 2008. Imaging via compressive sampling. *IEEE Signal Process. Mag*., **25**(2):14-20. [doi:10.1109/MSP.2007.914729]

Ross, A.A., Govindarajan, R., 2005. Feature level fusion of hand and face biometrics. Proc. SPIE, p.196-204. [doi:10.1117/12.606093]

Shi, W.Z., Zhu, C.Q., Tian, Y., *et al.*, 2005. Wavelet-based image fusion and quality assessment. *Int. J. Appl. Earth Observ. Geoinform.*, **6**(3-4):241-251. [doi:10.1016/j.jag.2004.10.010]

Smith, L.I., 2002. A Tutorial on Principal Components Analysis. Available from www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf.

Tropp, J., Gilbert, A.C., 2007. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans. Inform. Theory*, **53**(12):4655-4666. [doi:10.1109/TIT.2007.909108]

Wan, T., Qin, Z.C., 2011. An application of compressive sensing for image fusion. *Int. J. Comput. Math.*, **88**(18):3-9. [doi:10.1080/00207160.2011.598229]

Wang, R., Du, L.F., 2014. Infrared and visible image fusion based on random projection and sparse representation. *Int. J. Remote Sens.*, **35**(5):1640-1652. [doi:10.1080/01431161.2014.880819]

Wang, Y., Yang, J.F., Yin, W., *et al*., 2008. A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Image Sci.*, **1**(3):248-272. [doi:10.1137/080724265]

Xydeas, C.S., Petrović, V., 2000. Objective image fusion performance measure. *Electron. Lett.*, **36**(4):308-309. [doi:10.1049/el:20000267]

Yang, X.H., Jin, H.Y., Jiao, L.C., 2007. Adaptive image fusion algorithm for infrared and visible light images based on DT-CWT. *J. Infrared Millim. Waves*, **26**(6):419-424 (in Chinese).

Yang, Y., Han, C.Z., Kang, X., *et al.*, 2007. An overview on pixel-level image fusion in remote sensing. Proc. IEEE Int. Conf. on Automation and Logistics, p.2339-2344. [doi:10.1109/ICAL.2007.4338968]

Zheng, Y.Z., Qin, Z., 2009. Region-based image fusion method using bidimensional empirical mode decomposition. *J. Electron. Imag.*, **18**(1):013008. [doi:10.1117/1.3099703]