

# Simultaneous Image Fusion and Denoising based on Multi-Scale Transform and Sparse Representation

**Tahiatul Islam, Sheikh Md. Rabiul Islam**

Dept. of Electronics and Communication Engineering Khulna University of Engineering & Technology, Bangladesh.  
Email: tahiatul@gmail.com; robi@ece.kuet.ac.bd

**Xu Huang**

Faculty of ESTeM, University of Canberra, Australia,  
Email: Xu.Huang@canberra.edu.au

**Keng Liang Ou**

College of Oral Medicine, Taipei Medical University, Taiwan.  
Email: au,klou@tmu.edu.tw

**Abstract**—Multi-scale transform (MST) and sparse representation (SR) techniques are used in an image representation model. Image fusion is used especially in medical, military and remote sensing areas for high resolution vision. In this paper an image fusion technique based on shearlet transformation and sparse representation is proposed to overcome the natural defects of both MST and SR based methods. The proposed method is also used in different transformations and SR for comparison purposes. This research also investigate denoising techniques with additive white Gaussian noise into source images and perform threshold for de-noised into the proposed method. The image quality assessments for the fused image are used for the performance of proposed method and compared with others.

**Index Terms**—Image fusion; MST; SR; image denoising; Shearlet.

## I. INTRODUCTION

Image fusion is a process to take part for multi-view information into a new one image which contains high quality information. Researchers have developed algorithms based on image fusion containing high quality fused image with the help of two source images. This idea develops research area in remote sensing, medical image processing, and computer vision. Image fusion has three levels such as pixel, feature and decision level. In this paper we consider pixel level fusion [1], the fused image should keep the information and also remove the noise from the source images. The pixel level fusion is divided into spatial and transform domain [2]. The transform domain technique is used in image fusion based on multi-scale transforms. It performs on a number of different scales and orientations independently.

Recently, many multi-scale geometric analysis (MGA) tools have been proposed, such as Pyramid Transforms, the Discrete Wavelet Transform (DWT) [1] - [3], the Undecimated Wavelet Transform (UWT) [1], [3], the Dual-Tree Complex Wavelet Transform (DTCWT) [4], curvelet [5], bandelet [6], contourlet [7], shearlet [8], etc. for the representation of multidimensional data. Shearlet is a new multi-resolution approach provided higher directional sensitivity properties. The authors [8] have developed shearlet with a rich mathematical structure similar to wavelets but the decomposition procedures same as contourlet transform. The decomposition process of shearlet is followed with directional filter with shear matrix whereas contourlet transform is based on the Laplacian pyramid.

The problem of dictionary learning can be stated as follows [10]:

$$y = Dx, D \in \mathbb{R}^{m \times r}, x \in \mathbb{R}^{r \times n} \quad (1)$$

Where  $D$  is referred to as the dictionary matrix and  $x$  is the coefficient of matrix and both are unknown.  $n$  is the signal dimension and  $m$  is the dictionary size. The equation (1) can be solved in case of the solution to (1) is not unique when the number of dictionary element  $r \geq m$  (i.e. the dictionary is *overcomplete*). Consider the coefficient matrix  $x$  is sparse and can be observed  $y_i \in \mathbb{R}^{m \times r}$  is a sparse combination of the dictionary elements (i.e. columns of the dictionary matrix). This problem is known as *sparse coding*. The learning community have been many works on dictionary learning. Hillar and Sommer [11] the number of sample data in fit into dictionary is exponential in  $r$  for general case. Spielman et al. (2012) sparse coding fits the  $l_1$  norm based method for *undercomplete* dictionary  $r \leq m$ . In our paper we have considered  $l_1$  norm based method  $r \geq m$ . The sparse coding is mainly calculated the sparsest of  $x$ .

which contains the fewest nonzero elements. This coding method has shown stability and efficient of algorithm Yaghoobi et al. [10] .Yang and Li et al. [11] had used sliding window technique into Sparse coding and focused on fusion process more robust than noise and registration.

A hybrid model of image fusion using MST and SR based method is proposed here where the framework is compared with different MGA method and the result is analyzed. The main purpose of this paper is to use that hybrid model for shearlet transform and sparse representation and to use de-noising method to remove any noise subjected to source images.

This paper is organized firstly for giving the theoretical description of shearlet transform and sparse modeling then the framework for image fusion and comparison of the proposed method with existing method and finally the conclusion of it.

## II. SHEARLETS

The theory of composite wavelets provides an effective approach for combining geometry and multi-scale analysis by taking advantage of the classical theory of affine systems. In dimension n=2, the affine systems with composite dilations are defined as follows [8]

$$A_{AS(\psi)} = \left\{ \psi_{j,l,k}(x) = |\det A|^{\frac{j}{2}} \psi(S^l A^j x - k) : j, l \in \mathbb{Z}, k \in \mathbb{Z} \right\} \quad (2)$$

Where  $\psi \in l^2(\mathbb{R}^2)$  is called composite wavelet.  $A, S$  are both  $2 \times 2$  invertible matrix, and  $|\det S| = 1$ , the elements of this system are called composite wavelet if  $A_{AS(\psi)}$  forms a tight frame for

$$l^2(\mathbb{R}^2) \sum_{j,l,k} | \langle f, \psi_{j,l,k} \rangle |^2 = \|f\|^2$$

Let  $A$  denote the parabolic scaling matrix and  $S$  denote the shear matrix. For each  $a > 0$  and  $s \in R$ ,

$$A = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}, S = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$$

The size of frequency support of the shearlet is illustrated in Fig. 1 for some particular values of  $a$  and  $s$ . In [8], consider  $a = 4, s = 1$ ; where  $A = A_0$  is the anisotropic dilation matrix and  $S = S_0$  is the shear matrix, which are given by

$$A_0 = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}, S_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

For  $\forall \xi = (\xi_1 \xi_2) \in \widehat{\mathbb{R}}^2, \xi_1 \neq 0$ , let  $\widehat{\psi}^{(0)}(\xi)$  be given by

$$\widehat{\psi}^{(0)}(\xi) = \widehat{\psi}^{(0)}(\xi_1 \xi_2) = \widehat{\psi}_1(\xi_1) \widehat{\psi}_2\left(\frac{\xi_2}{\xi_1}\right)$$

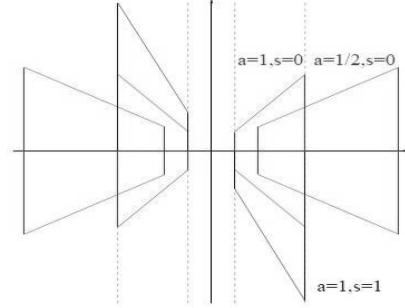


Fig.1. Frequency support of the shearlets  $\widehat{\psi}_{j,l,k}$  for different values of  $a$  and  $s$  [8].

Where  $\widehat{\psi}_1 \in C^\infty(\mathbb{R})$  is a wavelet, and  $\text{supp } \widehat{\psi}_1 \subset [-\frac{1}{2}, \frac{1}{16}] \cup [-\frac{1}{16}, \frac{1}{2}]$ ;  $\widehat{\psi}_2 \in C^\infty(\mathbb{R})$  and  $\text{supp } \widehat{\psi}_2 \subset [-1, 1]$ . This implies  $\widehat{\psi}^{(0)} \in C^\infty(\mathbb{R})$  and  $\widehat{\psi}^{(0)} \subset [-\frac{1}{2}, \frac{1}{2}]^2$ .

In addition, we assume that

$$\sum_{j \geq 0} |\widehat{\psi}_1(2^{-2j}\omega)|^2 = 1, |\omega| \geq 1/8 \quad (3)$$

and for  $\forall j \geq 0$

$$\sum_{l=-2^j}^{2^j-1} |\widehat{\psi}_2(2^j\omega - l)|^2 = 1, |\omega| \leq 1 \quad (4)$$

Equation (3) and (4) imply that

$$\sum_{j \geq 0} \sum_{l=-2^j}^{2^j-1} |\widehat{\psi}^{(0)}(\xi A_0^{-j} S_0^{-l})|^2 = \sum_{j \geq 0} \sum_{l=-2^j}^{2^j-1} |\widehat{\psi}_1(2^{-2j}\xi_1)|^2 |\widehat{\psi}_2(2^j \frac{\xi_2}{\xi_1} - l)|^2 = 1$$

For any  $(\xi_1, \xi_2) \in D_0$ , where  $D_0 = \{(\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : |\xi_1| \geq \frac{1}{8}, |\xi_2| \leq 1\}$ , the functions  $\{\widehat{\psi}^{(0)}(\xi A_0^{-j} S_0^{-l})\}$  form a tiling of  $D_0$ . This is illustrated in Fig. 2(a). This property described above implies that the collection.

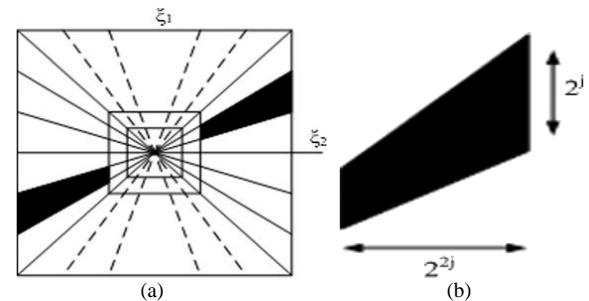


Fig.2. (a) The tiling of the frequency by the shearlets. (b) The size of the frequency support of a shearlet  $\widehat{\psi}_{j,l,k}$  [8].

$$\left\{ \psi_{j,l,k}^{(0)}(x) = 2^{\frac{3j}{2}} \widehat{\psi}^{(0)}(S_0^l A_0^j x - k) : j \geq 0, -2^j \leq l \leq 2^j - 1, k \in \mathbb{Z}^2 \right\}$$

is a Parseval frame for  $L^2 = (D_0)^V \{f \in L^2(\mathbb{R}^2) : \text{supp } f \in D_0\}$  based on the support of  $\hat{\psi}_1$  and  $\hat{\psi}_2$  can easily be observe that the functions  $\psi_{j,l,k}$  have frequency support,

$$\text{supp } \hat{\psi}_{j,l,k}^{(0)} \subset \left\{ (\xi_1, \xi_2) : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{-2j-4}, 2^{2j-1}], \left| \frac{\xi_2}{\xi_1} + l2^j \right| \leq 2^{-j} \right\}$$

That is, each element  $\hat{\psi}_{j,l,k}$  is supported on a pair of trapezoids, of approximate size  $2^{2j} \times 2^j$ , oriented along lines of slope  $l2^j$  as shown in Figure 2 (b).

Similarly we can construct a Parseval frame for  $L^2(D_1)^V$ , where  $D_1$  is the vertical cone,

$$D_1 = \left\{ (\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_2| \geq 1/8, \left| \frac{\xi_2}{\xi_1} \right| \leq 1 \right\}$$

Let  $A = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$  and  $S = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$   
and  $\hat{\psi}^{(1)}$  be given by

$$\hat{\psi}^{(1)}(\xi) = \hat{\psi}^{(1)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\hat{\psi}_2\left(\frac{\xi_1}{\xi_2}\right)$$

where  $\hat{\psi}_1$  and  $\hat{\psi}_2$  are defined above. Then the collection [8]

$$\hat{\psi}_{j,l,k}^{(1)} = 2^{\frac{3j}{2}} \hat{\psi}^{(1)}(S_1^l A_1^j x - k) : j \geq 0, -2^j \leq 2^j, k \in \mathbb{Z}^2$$

Is a Parseval frame for  $L^2(D_1)^V$ . Finally let  $\hat{\psi} \in C_0^\infty(\mathbb{R}^2)$  be chosen to satisfy

$$G(\xi) = |\hat{\psi}(\xi)|^2 + R_j + P_j = 1$$

Where

$$R_j = \sum_{l \geq 0} \sum_{i=-2^j}^{2^j-1} |\hat{\psi}^{(0)}(\xi S_0^l A_0^j x)|^2 \chi_{D_0}(\xi)$$

$$P_j = \sum_{l \geq 0} \sum_{i=-2^j}^{2^j-1} |\hat{\psi}^{(1)}(\xi S_1^l A_1^j x)|^2 \chi_{D_1}(\xi) = 1$$

for  $\xi \in \mathbb{R}^2$

where  $\chi_D$  denotes the indicator function of the set  $D$ . This implies that  $\hat{\psi} \subset \left[-\frac{1}{8}, \frac{1}{8}\right]^2$  with  $|\hat{\psi}| = 1$  for  $\xi \subset \left[-\frac{1}{8}, \frac{1}{8}\right]^2$  and the set  $\{\psi(x - k) : k \in \mathbb{Z}^2\}$  is a Parseval frame for  $L^2\left([- \frac{1}{16}, \frac{1}{16}]^2\right)^V$ . It is observed that, by the properties of  $\psi^{(d)}$ ,  $d = 0, 1$ , it follows that the function  $G(\xi) = (\xi_1, \xi_2)$  is continuous and regular along the lines  $\frac{\xi_2}{\xi_1} = \pm 1$  (as well as for any other  $\xi \in \mathbb{R}^2$ ) [8].

### III. BASIC THEORY OF SIGNAL SPARSE REPRESENTATION (SR)

SR is based on linear combination of a prototype signal from dictionary [12]. The following notation [12], considered  $f \in \mathbb{R}^n$  be  $n$  real valued samples signal. The sparse representation theory the dictionary is formed on  $D \in \mathbb{R}^{n \times m}$ , which contains  $m$  prototype signals. This dictionary  $D$  which determines the signal is sparse. In the dictionary  $D$ , there exists a linear combination of prototype signal and it shows  $\forall x \in f, \exists s \in \mathbb{R}^T$  such that  $x \approx Ds$ , where  $s$  is sparse coefficients of  $x$  in  $D$ . It usually assumes that  $m > n$ , implying that the dictionary  $D$  is redundant and it follows Restricted Isometric Property (RIP) [12]. This property solves the problem of reconstruction of signal with optimization problem to find the smallest possible number of nonzero components of  $s$ :

$$\min_s \|s\|_0 \text{ subject to } \|Ds - x\| < \varepsilon \quad (5)$$

This is called  $l_0$ -norm and  $\|s\|_0$  denoting the number of nonzero coefficients in sparses. This is an NP-hard problem. This problem can only be solved by systematically testing based on combinations of columns  $D$ . This problem is solved by greedy algorithms such as the basis pursuits (BP), matching pursuit (MP) and the orthogonal MP (OMP) algorithms [13].

### IV. SPARSE REPRESENTATION FOR IMAGE FUSION

The sparse representation (SR) handles an image in globally. But it cannot handle image fusion because it depends on the local information of source images. The source images are divided into small patches with fixed dictionary  $D$  to solve this problem. In addition, a sliding window technique and shift invariant are also important to image fusion for the sparse representation.

We consider that a source image  $I$  is divided into  $j$  small patches by size  $n \times n$  as shown in Figure 3. It can be represented into lexicographically orders as a vector  $v$ . Then,  $V^j$  can express as

$$v = \sum_{t=1}^T S^j(t)d_t \quad (6)$$

where  $j$  is the number of image patches and  $d_t$  is an atoms/prototype from  $D$  and  $D = [d_1 \dots d_t \dots d_T]$ , which contains  $T$  prototypr vectors.  $S^j = [s^1(1) \dots s^j(t) \dots s^j(T)]$  is the sparse representation.

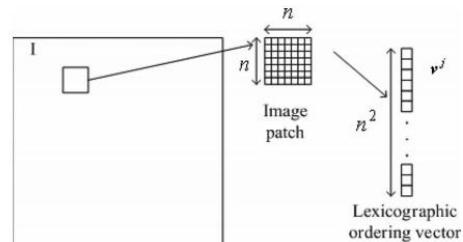


Fig.3. A source image patches and its lexicographic ordering vector.

The image patches of  $I$  are now rebuild into one matrix  $v$ . Then,  $v$  can be expressed as

$$v = [d_1 \dots d_t \dots d_T] \begin{bmatrix} S^1(1) & S^2(1) & \dots & S^j(1) \\ S^1(2) & S^2(2) & \dots & S^j(2) \\ \vdots & \vdots & \ddots & \vdots \\ S^1(T) & S^2(T) & \dots & S^j(T) \end{bmatrix} \quad (7)$$

Then the equation (7) can be expressed as

$$v = DS$$

where  $S$  is a sparse matrix. Figure 4 represents the basic approach to fuse the image by sparse representation.

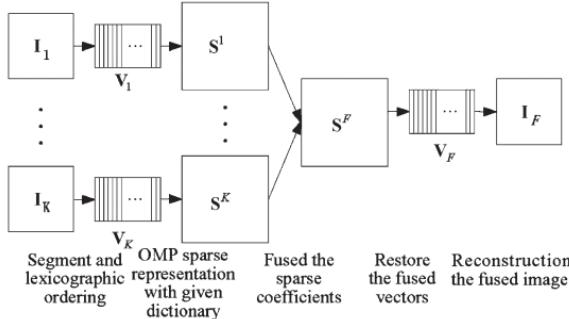


Fig.4. Proposed SR-based fusion method.

## V. IMAGE DENOISING

Images are subject to noise which are captured and transferred from one destination to another. Noise can greatly affect the fusion result. So to make the fusion more effective this paper introduces thresholding to the respective coefficient values after taking the MST decomposition. We consider BayesShrink threshold which performs better than SureShrink in literature in terms of mean square error (MSE).

The BayesShrink threshold is given by [14]:

$$\hat{T}(\hat{\sigma}_I) = \frac{\hat{\sigma}_n}{\hat{\sigma}_I} \quad (8)$$

where  $\sigma_n$  and  $\sigma_I$  are standard deviation of noise and image source respectively. Now we can find out the noise variance  $\sigma_n^2$  from the sub band details. We can also find out the median estimator on the 1-D sub band coefficients by

$$\hat{\sigma}_n = \text{median}(|\text{Details}|)/0.6745 \quad (9)$$

The noise model of image  $x$  is considered as  $x = I + n$  and source image ( $I$ ) and source noise ( $n$ ) are assumed to be independent. Therefore the average noise (variance) and signal power can be calculated as,

$$\sigma_x^2 = \sigma_I^2 + \sigma_n^2 \quad (10)$$

where  $\sigma_x^2$  is the variance of the observed signal  $\sigma_n^2$  is the noise density for source image. The output noise power of source image  $\sigma_I^2$  is

$$\hat{\sigma}_I = \sqrt{\max((\hat{\sigma}_s^2 - \hat{\sigma}_n^2), 0)} \quad (11)$$

## VI. PROPOSED METHOD

The proposed method based on Shearlet transform (ST) and sparse representation (SR) into image fusions are described in section II, III & IV. The main researcher work of proposed method is to extract the high-frequency information from image by MST method. SR is applied to prevent and decrease contrast and spectral information distortion for the fused image. The source images have some noises detected and apply a threshold is applied a given in Equation (7). The block diagram of ST with SR proposed method is shown in Fig. 5.

The proposed image fusion methodology contains the following five steps which are described below:

- A. *MST Decomposition*: Apply MST on the two source images  $\{I_A, I_B\}$  to obtain their low-pass bands  $\{L_A, L_B\}$  and high-pass bands  $\{H_A, H_B\}$ .
- B. *Thresholding*: Perform thresholding as obtained from Equation (7) on the low-pass and high-pass to remove the unnecessary coefficients from the decompositon.
- C. *Low-pass fusion*:
  - i. The image  $I$  is divided by applying sliding window technique to  $\{I_A, I_B\}$  into image patches of size  $\sqrt{n} \times \sqrt{n}$  from upper left to lower right with a step length of pixels. We have also considered that there are  $T$  patches denoted as  $\{P_A^i\}_{i=1}^T$  and  $\{P_B^i\}_{i=1}^T$  in  $L_A$  and  $L_B$  respectively.
  - ii. Rearrange  $\{P_A^i, P_B^i\}$  into column vectors and also rearrange  $\{\hat{V}_A^i, \hat{V}_B^i\}$  for each iteration by

$$\begin{aligned} \hat{V}_A^i &= V_A^i - \bar{V}_A^i \cdot 1 \\ \hat{V}_B^i &= V_B^i - \bar{V}_B^i \cdot 1 \end{aligned}$$

where 1 denotes one dimensional vector  $asn \times 1$  and  $\bar{V}_A^i$  and  $\bar{V}_B^i$  are average values of all elements in  $V_A^i$  and  $V_B^i$  respectively.

- iii. The sparse coefficients vectors  $\{\alpha_A^i, \alpha_B^i\}$  of  $\{\hat{V}_A^i, \hat{V}_B^i\}$  is calculated using OMP algorithm [13] by

$$\begin{aligned} \alpha_A^i &= \text{argmin} \|\alpha\|_0 \text{ st } \|\hat{V}_A^i - D\alpha\|_2 < \epsilon \\ \alpha_B^i &= \text{argmin} \|\alpha\|_0 \text{ st } \|\hat{V}_B^i - D\alpha\|_2 < \epsilon \end{aligned}$$

Where  $D$  is the learned dictionary and  $\epsilon$  is tolerance of sparse value.

- iv. Merge  $\alpha_A^i$  and  $\alpha_B^i$  with max-L1 rule to obtained the fused sparse vectors

$$\alpha_F^i = \begin{cases} \alpha_A^i & \text{if } \|\alpha_A^i\|_1 > \|\alpha_B^i\|_1 \\ \alpha_B^i & \text{otherwise} \end{cases}$$

The fused result of  $V_A^i$  and  $V_B^i$  is calculated by  
 $V_F^i = D\alpha_F^i + \bar{V}_F^i \cdot 1$

Where the merged mean value  $\bar{V}_F^i$  is obtained by  $\bar{V}_F^i = \begin{cases} \bar{V}_A^i & \text{if } \alpha_F^i = \alpha_A^i \\ \bar{V}_B^i & \text{otherwise} \end{cases}$

- v. The above iterative procedures for all the source image patch in  $\{P_A^i\}_{i=1}^T$  and  $\{P_B^i\}_{i=1}^T$  to obtained all the fused vectors  $\{P_F^i\}_{i=1}^T$  and  $\{V_F^i\}_{i=1}^T$ .

vi. At the final steps the low-pass fused results denoted  $L_F$ . For each  $V_F^i$  reshape it into a patch  $P_F^i$  and then plug  $P_F^i$  into its original position in  $L_F$ . As patches are overlapped, each pixel's value in  $L_F$  is averaged over its accumulation times.

D. *High-pass fusion:* Apply threshold rule using Equation (7) for filter to ensure fused image contains source image. And then merge  $H_A$  and  $H_B$  to obtained  $H_F$  with the popular ‘max-absolute’ rule each coefficients of image A and B.

E. *MST reconstruction:* Perform the corresponting inverse MST over  $L_F$  and  $H_F$  to reconstruct the final fused image  $I_F$ .

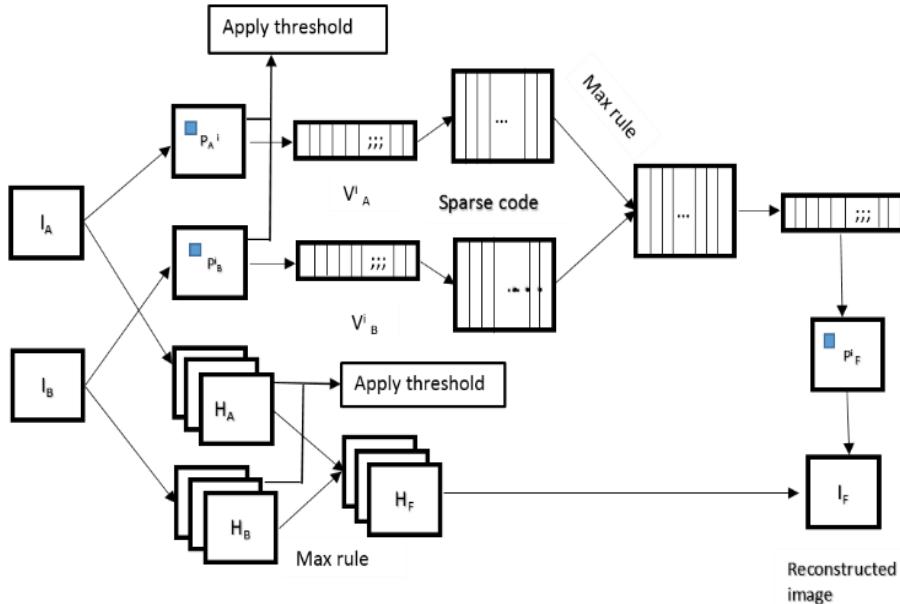


Fig.5. Block diagram of proposed method.

## VII. EXPERIMENTS AND ANALYSIS

The source image pairs are taken from database [15 as shown in Figure 6.] and grouped of these images to apply proposed ST-SR method for image fusion and compare the results with other MST methods.

For dictionary process in image sparse presentation, we assign randomly values in dictionary  $D$  of size  $64 \times 256$  from the image patches of training dataset image then sparse coding is done to get the sparse matrix of signal then K-SVD algorithm [11] has been run to update the dictionary. In simulation, we have estimated 100,000 of training data and  $8 \times 8$  patches, and these patches are randomly sampled into image. In simulation, the dictionary size was set to 256 and the iteration number of K-SVD is fixed to 180.

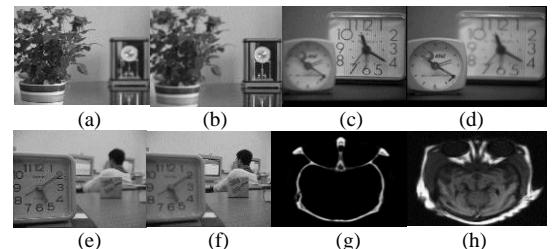


Fig.6. Source images pair for experimental analysis[14].

For comparative studies, we have used different MST method and SR such as discrete wavelet transform with sparse representation (DWT-SR), dual-tree complex wavelet transform(DTCWT) and sparse representation (DTCWT-SR), curvelet transform and sparse representation (CT-SR), nonsubsampled contourlet transform (NSCT) and sparse representation (NSCT-SR), non-subsampled shearlet transform (NSST) with sparse representation (NSST-SR).

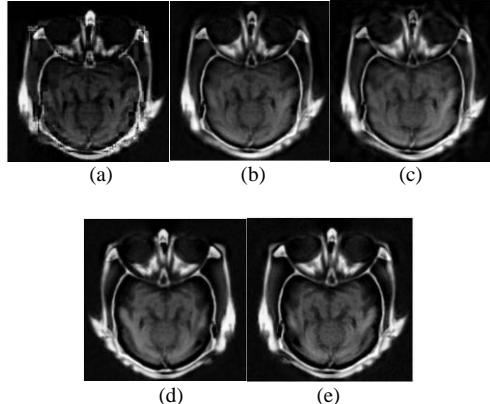
Table 1. Performance measurements of different image fusion method.

	Methods	Fig.4 (a)(b)	Fig.4 (c)(d)	Fig.4 (e)(f)	Fig.4 (g)(h)
SD	DWT-SR	52.2803	51.1927	47.2138	51.7858
	DTCWT-SR	51.8128	50.8592	46.9703	53.5718
	CT-SR	51.9511	50.8515	46.8730	54.4996
	NSCT-SR	52.1892	51.1708	47.2209	56.8796
	<b>NSST-SR</b>	<b>52.9218</b>	<b>51.9778</b>	<b>47.9486</b>	<b>56.9551</b>
MI	DWT-SR	6.4716	6.9422	7.0805	2.2327
	DTCWT-SR	6.9836	7.0816	7.4136	2.0315
	CT-SR	6.6547	7.0543	6.9303	1.9097
	NSCT-SR	6.9077	7.6748	7.5160	3.0118
	<b>NSST-SR</b>	<b>6.9857</b>	<b>7.9168</b>	<b>7.9504</b>	<b>3.1877</b>
EN	DWT-SR	7.4870	7.3228	7.1110	6.1926
	DTCWT-SR	7.4448	7.3318	7.0817	6.7317
	CT-SR	7.4642	7.3946	7.0999	6.8621
	NSCT-SR	7.4608	7.3292	7.0962	6.5892
	<b>NSST-SR</b>	<b>7.6676</b>	<b>7.5513</b>	<b>7.4051</b>	<b>6.9054</b>
SSIM	DWT-SR	0.8810	0.8907	0.9022	0.5449
	DTCWT-SR	0.8865	0.8938	0.9078	0.7004
	CT-SR	0.8862	0.8941	0.9070	0.6811
	NSCT-SR	0.8918	0.9015	0.9112	0.7637
	<b>NSST-SR</b>	<b>0.8994</b>	<b>0.9098</b>	<b>0.9195</b>	<b>0.7692</b>
$Q^{AB/F}$	DWT-SR	0.6946	0.6810	0.7038	0.5773
	DTCWT-SR	0.7166	0.7102	0.7277	0.6334
	CT-SR	0.7023	0.6994	0.7069	0.5898
	NSCT-SR	0.7132	0.7216	0.7303	0.7368
	<b>NSST-SR</b>	<b>0.7195</b>	<b>0.7271</b>	<b>0.7358</b>	<b>0.7394</b>
$Q_G$	DWT-SR	0.6874	0.6912	0.7171	0.5645
	DTCWT-SR	0.7071	0.7192	0.7368	0.6159
	CT-SR	0.6867	0.7087	0.7192	0.5713
	NSCT-SR	0.7025	0.7320	0.7400	0.7329
	<b>NSST-SR</b>	<b>0.7096</b>	<b>0.7371</b>	<b>0.7473</b>	<b>0.7394</b>

In this research work, five popular metrics have been used for evaluating the quality of a fused image [16], which are Standard deviation (SD), Mutual information (MI), Entropy (EN), Standard similarity index metric (SSIM),  $Q^{AB/F}$ , the gradient based image fusion metric  $Q_G$  represents in Table I. In Table I shows the proposed method such as NSST-SR is achieved the highest results in bold face mark for SD, MI, EN SSIM,  $Q^{AB/F}$ ,  $Q_G$  than DWT-SR, DTCWT-SR, CT-SR, NSCT-SR. Also the fused image of NSST-SR shows good quality visual fused image in Figure 7 (e) for comparison with them.

Table 2. Performance measurements with different noise level for different image fusion method.

Sigma( $\sigma$ )	Methods	PSNR
10	DWT-SR	31.4988
	DTCWT-SR	31.9889
	CT-SR	31.8623
	NSCT-SR	31.8978
	NSST-SR	36.2864
20	DWT-SR	29.3442
	DTCWT-SR	29.6937
	CT-SR	29.5617
	NSCT-SR	29.5967
	NSST-SR	33.8054
30	DWT-SR	28.5474
	DTCWT-SR	28.7777
	CT-SR	28.7332
	NSCT-SR	28.7283
	NSST-SR	32.7254
40	DWT-SR	28.1573
	DTCWT-SR	28.3370
	CT-SR	28.3004
	NSCT-SR	28.2909
	NSST-SR	32.1793
50	DWT-SR	27.9320
	DTCWT-SR	28.0779
	CT-SR	28.0589
	NSCT-SR	28.0208
	NSST-SR	31.7270

Fig.7. Fused image of (a) DWT-SR, (b) DTCWT-SR, (c) CT-SR, (d) NSCT-SR, (e) NSST-SR, for  $\sigma = 0$ .

Before we have discussed that if there are noises in the image fusion databases, what will happen in the proposed method? For experimental analysis, we added different noise level i.e., different standard deviation ( $\sigma$ ) for AWGN and applied threshold of Equation (7) after MST

to remove the noise and get good quality de-noised image. The de-noised image is then applied for sparse presentation. For comparative study, we have used different MST method and SR as shown in Table II. Table II represents the PSNR value for different noise level ( $\sigma$ ) added to the image shown in Figure. 6(g) & (h) as an example. We have applied this technique into all of 40 images and taken the results but because of limitation of paper area we can't give all the results. The Table II shows the value of PSNR of different MST method with SR. All of these methods, NSST-SR method give higher PSNR value for all case of noise level than DWT-SR, DTCWT-SR, CT-SR, NSCT-SR and also shown in Figure 8.

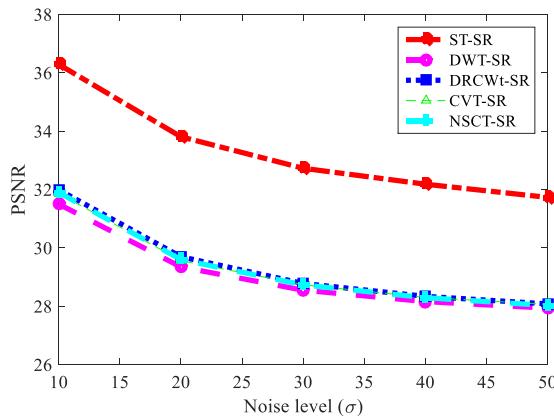


Fig.8. PSNR variation for different noise value of  $\sigma$  for multifocus image of image pair shown in Figure 4(g) & (h).

## VIII. CONCLUSION

This paper work proposes a method for image fusion using MST and sparse representation. It gives us the result which overcomes the disadvantage introduced in MST and SR method individually. Training time for Dictionary in SR is relatively high. We use a dataset consisting different type of images and which contain about 40 images. For thresholding the image we have used bayesShrik thresholding which take the threshold value in different direction of image under certain transformation domain, which gives good threshold and results in a fused image which is noiseless in great measure. In the future, this model can be enhanced and cope with different efficient transformation domain which can take the geometry of image in better way. So it will be interesting to see how advancement is made in this approach.

## REFERENCES

- [1] Z. Zhang and R. S. Blum, "A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application," in *Proceedings of the IEEE*, vol. 87, no. 8, pp. 1315-1326, Aug 1999.
- [2] H. Li, B. S. Manjunath, S. K. Mitra, "Multisensor Image Fusion Using the Wavelet Transform," *Graphical Models and Image Processing*, Volume 57, Issue 3, pp. 235-245, 1995.
- [3] V. S. Petrovic and C. S. Xydeas, "Gradient-based multiresolution image fusion," *IEEE Transaction Image Processing*, Volume. 13, No. 2, pp. 228–237, Feb. 2004.
- [4] J. J. Lewis, J. Robert, O'Callaghan, S. G. Nikolov, D. R. Bull, Nishan Canagarajah, "Pixel- and region-based image fusion with complex wavelets," *Information Fusion*, Volume 8, Issue 2, , pp. 119-130, April 2007
- [5] F. Nencini, A. Garzelli, S. Baronti, L. Alparone, "Remote sensing image fusion using the curvelet transform," *Information Fusion*, Volume 8, Issue 2, pp. 143-156, April 2007.
- [6] Villegas, O.O.V., De Jesus Ochoa Dominguez, H., Sanche, V.G.C., "A Comparison of the Bandelet, Wavelet and Contourlet Transforms for Image Denoising," *Seventh Mexican International Conference on Artificial Intelligence, MICAI 2008*, pp. 207–212, 2008.
- [7] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," in *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091-2106, Dec. 2005.
- [8] W. Q. Lim, "The Discrete Shearlet Transform: A New Directional Transform and Compactly Supported Shearlet Frames," *IEEE Transactions on Image Processing*, Volume. 19, No. 5, pp. 1166-1180, 2010.
- [9] Yu Liu, Shuping Liu, Zengfu Wang, "A general framework for image fusion based on multi-scale transform and sparse representation," *Information Fusion*, Volume 24, July 2015, pp. 147-164,
- [10] A. Agarwal, A. Anandkumar, P. Jain, P. Netrapalli, R. Tandon "Learning Sparsely Used Overcomplete Dictionaries," *JMLR: Workshop and Conference Proceedings*, vol 35, pp. 1–15, 2014
- [11] Christopher J Hillar and Friedrich T Sommer. Ramsey theory reveals the conditions when sparsecoding on subsampled data is unique. *arXiv preprint arXiv:1106.3616*, 2011
- [12] D. A. Spielman, H. Wang, and J. Wright, "Exact recovery of sparsely-used dictionaries," *Journal of Machine Learning Research* pp.1–35, 2012.
- [13] M. Yaghoobi, T. Blumsenath and M. E. Davies, "Dictionary Learning for Sparse Approximations With the Majorization Method," *IEEE Transactions on Signal Processing*, Volume. 57, No. 6, pp. 2178-2191, 2009.
- [14] B. Yang and S. Li, "Multifocus Image Fusion and Restoration with Sparse Representation," in *IEEE Transactions on Instrumentation and Measurement*, Volume 59, No. 4, pp. 884-892, 2010
- [15] M. Yaghoobi, T. Blumsenath and M. E. Davies, "Dictionary Learning for Sparse Approximations With the Majorization Method," *IEEE Transactions on Signal Processing*, Volume. 57, No. 6, pp. 2178-2191, 2009.
- [16] B. Yang and S. Li, "Multifocus Image Fusion and Restoration with Sparse Representation," in *IEEE Transactions on Instrumentation and Measurement*, Volume 59, No. 4, pp. 884-892, 2010.
- [17] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transaction Signal Processing*, Volume 54, No. 11, pp. 4311-4322, 2006.
- [18] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, Volume 53, No. 12, pp. 4655 – 4666, 2007.

- [19] D. L. Donoho, "De-noising by soft thresholding," *IEEE Transaction Information Theory*, vol. 41, no. 3 pp. 613-627, May 1995.
- [20] Image fusion. Some source images (multi-focus, multi-modal) for image fusion: <http://home.ustc.edu.cn/~liuyu1/>
- [21] P. Jagalingam, Arkal Vittal Hegde, "A Review of Quality Metrics for Fused Image," *Aquatic Procedia*, Volume 4, 2015, pp. 133-142.

### Authors' Profiles



**Tahiatul Islam** received the bachelor degree from Khulna University of Engineering and Technology. He is currently working as a Software Engineer in Bangladesh. His main research area includes Image processesing and Machine lerning.



**Sheikh Md. Rabiul Islam** received the B.Sc.in Engg. (ECE) from Khulna University, Khulna, Bangladesh in December 2003, and M.Sc. in Telecommunication Engineering from the University of Trento, Italy, in October 2009 and Ph.D. from University of Canberra, Australia, in 2015. He joined as a Lecturer in the department of Electronics and Communication Engineering of Khulna University of Engineering & Technology, Khulna, in 2004, where he is joined an Assistant Professor in the same department in the effect of 2008.Also he joined Associate Professor in the same department. He has published 3 book chapters, 26 Journal and 15 International conferences. His research interests include biomedical signal & image processing, VLSI for signal processing and Wireless Communications. He is an IEEE member.



**Professor (Dr) Xu Huang** has received the B.E. and M.E. degrees and Ph.D. in Electrical Engineering and Optical Engineering prior to 1989 and the second Ph.D. in Experimental Physics in the University of New South Wales, Australia in 1992. He has earned the Graduate Certificate in Higher Education in 2004 at the University of Canberra, Australia. He has been working on the areas of the

telecommunications, cognitive radio, networking engineering, wireless communications, optical communications, and digital signal processing more than 30 years. Currently he is the Professor at the Faculty of Education Science Technology & Mathematics. He has been a senior member of IEEE in Electronics and in Computer Society since 1989 and a Fellow of Institution of Engineering Australian (FIEAust), Chartered Professional Engineering (CPEng), a Member of Australian Institute of Physics. He is a member of the Executive Committee of the Australian and New Zealand Association for Engineering Education, a member of Committee of the Institution of Engineering Australia at Canberra Branch. Professor Huang is Committee Panel Member for various IEEE International Conferences such as IEEE IC3PP, IEEE NSS, etc. and he has published about two hundred papers in high level of the IEEE and other Journals and international conference; he has been awarded 9 patents in Australia.



**Professor Keng-Liang Ou** obtained Ph.D. degree from Graduate Institute of Mechanical Engineering, National Chiao Tung University, Taiwan. He came to Taipei Medical University, Taiwan to pursue cutting edge research for biomaterials and currently holds the position of elected Dean of College of Oral Medicine. He is also in charge of Graduate Institute of Biomedical Materials and Tissue Engineering, Research Center for Biomedical Implants and Microsurgery Devices and Research Center for Biomedical Devices and Prototyping Production. In addition to institutional appointment, Prof. Ou also serves as the President of Institute of Plasma Engineering Taiwan, the leader of the Taiwan society for metal heat treatment and the Head of Taiwan Oral Biomedical Engineering Association. Professor Ou devotes energy to novel research in the fields of biomaterials, bioengineering, biosensing and bioimaging, as well as translation medicine. In addition, he has extensive collaborations with industry and has played a major role in developing medical devices for health professionals worldwide. He is the leader and organizer for the biomedical product design, production, manufacturing, testing, legalization and marketing planning, with supports from team of scientists and researchers with different expertise. These efforts were well recognized by the community and Professore Ou was awarded with the 49<sup>th</sup> Ten Outstanding Young Persons of Taiwan at the year of 2011.

**How to cite this paper:** Tahiatul Islam, Sheikh Md. Rabiul Islam, Xu Huang, Keng Liang Ou, "Simultaneous Image Fusion and Denoising based on Multi-Scale Transform and Sparse Representation", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.9, No.6, pp.37-44, 2017.DOI: 10.5815/ijigsp.2017.06.05