

Fusion of Multi Focus Images Using Multi Scale Transform and Sparse Representation

MEGHANA T J

M.Tech, Digital Communication Engineering
M S Ramaiah Institute Of Technology
Bangalore,India

Prof.PARIMALA P

Associate Professor,Telecommunication Engineering
M S Ramaiah Institute Of Technology
Bangalore,India

Abstract—Multifocus image fusion is the process of integrating multiple source images having different focus of the same scene resulting in a single image,retaining the important information from each source image.In this paper we present a fusion method by combining Multiscale transforms(MST) method with Sparse Representation(SR).Source images are decomposed using MST to get low frequency components and high frequency components of source images.High frequency components are combined using absolute values of coefficients and low frequency components are merged using SR based fusion approach.The resultant fused image is obtained by applying inverse MST.Here we use two popular MSTs ,Discrete wavelet transform(DWT) ,Laplacian transform(LP) with levels ranging from one to four.The proposed methods are evaluated by various metrics such as Peak Signal to Noise Ratio(PSNR),Mean Absolute Error(MAE),Signal to Noise Ratio(SNR),Structural similarity Index(SSIM).

Keywords—Multifocus image fusion, Laplacian Pyramid, Discrete Wavelet Transform,Sparse Representation

I. INTRODUCTION

Optical lenses used in image sensing cameras have limited depth of focus,because of which it is not possible to get an image which contain all relevant objects in focus [1].However,for accurate interpretation and analysis of the scene,it is required to obtain images with every object in focus.A scene may contain many objects at varying distance.Due to the limited depth of camera,effective focusing on all objects in a scene is not likely possible.Which means certain images might be focused and certain images might be de-focused.

Image fusion has various applications like remote sensing, satellite imaging, medical imaging,military applications etc.Also the advantages include improved reliability,decreased uncertainty and storage expenses since a single image can store information from several images by fusion process[3].

Spatial domain and transform domain are two different methods used for image fusion.Spatial domain approach often produce undesirable effects namely contrast reduction and blurry edges.To reduce these effects transform domain approach is used [2]. Multi scale transform methods include the following three steps :First,to decompose image to get coefficients.Second,to merge the coefficients using fusion rule and lastly apply inverse transform to get fused image. Here we use two popular MSTs ,Discrete wavelet transform(DWT)[4],Laplacian transform(LP)[5].

In addition to selection of transform method,fusion rule in either high pass or low pass band also has a great impact on fused results.Having all these points in view,an multifocus image fusion algorithm is proposed.Initially MST is applied to decompose source images into low frequency and high frequency components. Since high frequency components contain information related to the structure of an image applying SR based rule may cause information to be lost hence here we apply SR based rule on low frequency components.Max absolute rule is applied on high frequency components in which we select components having maximum intensity value.Both the methods are evaluated using various parameters like PSNR,MAE,SNR,SSIM[6].This paper is organised as follows:In section 2 the implementation method including theory on Discrete wavelet transform,Laplacian Pyramid and Sparse Representation and proposed fusion scheme is briefed.The parameters used in evaluating the fusion algorithm is explained in Section 3.Experimental results are discussed in section 4 and section 5 concludes the paper.

II. IMPLEMENTING METHOD

1. Creating Image Dataset

Here we create an image dataset consisting of gray scale images.These images are taken from different websites.

2. Pre Processing

For an image taken from the data set,we apply gamma correction rule to vary the contrast of the choosen image,masking to smoothen the noise in the selcted image

and also apply resizing since the images to be fused should be of the same size.

3. MST and SR theories

3.1 Discrete Wavelet Transform:

Wavelet transform decomposes the image into its low frequency and high frequency bands. High frequency bands contain salient features such as edges or lines and low frequency bands contain directional information[6].

For implementation of DWT, 'Subband coding' is preferred. Subband coding uses low pass filter (H_0) and high pass filter (H_1) firstly, subband coding is applied on rows then on columns as shown in fig.1. After first level of decomposition, there will be four frequency bands, namely Low-Low (LL), Low-High (LH), High-Low (HL) and High-High (HH). The next level decomposition is just applied to the LL band of the current decomposition stage, which forms a recursive decomposition procedure. Thus, N-level decomposition will finally have $3N+1$ different frequency bands, which include $3N$ high frequency bands and just one LL frequency band.[7]

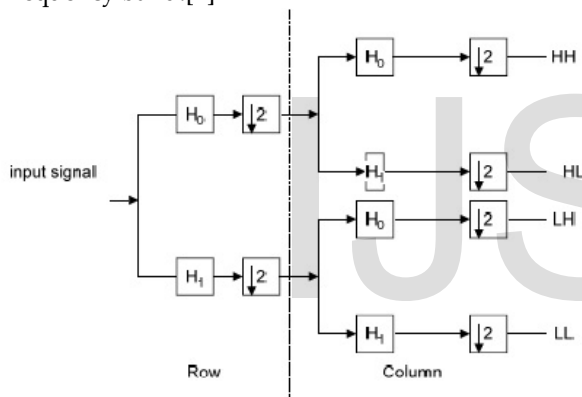


Fig 1: Filter bank structure of the DWT Analysis

For Image reconstruction inverse method of the subband coding is followed. The filter bank structure is shown below in fig.2

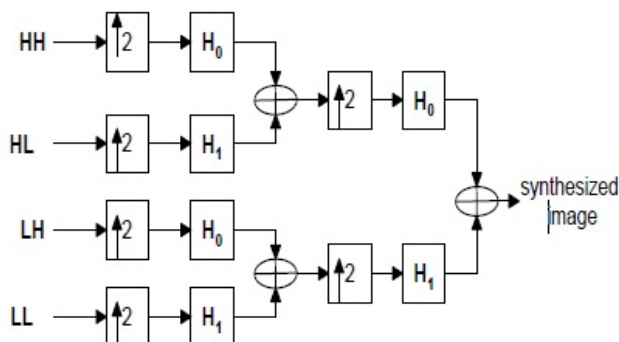


Fig 2: synthesis filter bank for IDWT

In figure 2, upsampling or interpolation is carried out and a pair of low pass (H_0) and high pass (H_1) filters are used which are same as that of in analysis. Matlab has provided the instruction for transform and inverse transform. The low

pass filtering coefficients are called 'approximation' and high pass filtering coefficients are called 'detail'.

3.2 Laplacian Pyramid:

A pyramid decomposition fusion consists of a number of images at different scales which together represent the original image. In general, every pyramid transform consists of three major phases:[6]

i. Decomposition

In Decomposition process a pyramid is generated in succession at each level of the fusion. The depth of fusion or number of levels of fusion is predefined. Decomposition phase basically consists of the following steps. These steps are performed 'l' number of times, 'l' being the number of levels to which the fusion will be performed.

1. The input images are first passed through a low pass filter with which these images are convolved/filtered.
2. After that the pyramid is generated from the convolved/filtered images.
3. The input images are then decimated to half their size, which would act as the input image matrices for the next level of decomposition.

ii. Formation of the initial image for re-composition

The input images are merged after the decomposition process. This resultant image would be used as the initial input to the re-composition process. The finally decimated input images are worked upon either by averaging the decimated input images, selecting the minimum decimated input image or selecting the maximum decimated input image.

iii. Re-composition

In the recombination process, the resultant image is finally created from the pyramids formed at each level of decomposition. These steps are performed 'p' number of times as in the decomposition process.

1. The input image is undecimated to the level of recombination
2. The transpose of the filter vector used in the decomposition stage is applied to convolve/filter the undecimated matrix.
3. The filtered matrix is then combined, by the process of pixel intensity value addition, with the pyramid formed at the respective level of decomposition.
4. This newly formed image matrix acts as the input to the next level of recombination.

The combined image at the final level of recombination will be the resultant fused image.

3.3 Sparse Representation:

Sparse representations have recently drawn much interest in vision, signal and image processing[8,9].

It represents a signal by a linear combination of "few" atoms from an overcomplete dictionary. Overcompleteness and sparsity are major characteristics of sparse representation. Overcompleteness means that the number of atoms in a dictionary is greater than the signal dimension. In sparse representation modeling, an input signal $y \in R^n$ can be represented by

$$y = D\alpha \quad (1)$$

where the vector $\alpha \in R^k$ contains the representation coefficients of the signal y and the dictionary matrix $D \in R^{n \times k}$ ($K > n$) contains K prototype signals referred as atoms. The sparsest prototype signals referred as atoms. The sparsest solution with the fewest number of nonzero coefficients can be found by solving the following optimization problem.

$$\min \|\alpha\|_0 \text{ s.t. } \|y - D\alpha\|_2 \leq \varepsilon \quad (2)$$

where $\|\cdot\|_0$ is the l_0 semi-norm that counts the number of nonzero entries in a vector and $\varepsilon > 0$ is the error tolerance. The above optimization problem is a NP-hard problem and can be solved approximately by greedy approaches such as Orthogonal Matching Pursuit (OMP) algorithm [14].

Constructing a proper dictionary is a key issue in sparse representation. Generally, there are two ways to obtain a dictionary. The first category is using analytical models such as Discrete Cosine Transform (DCT) and Curvelet Transform

(CVT). The second category is applying machine learning technique to obtain the dictionary from a large number of training image patches. The dictionary learning model can be represented by,

$$\min_{D, \{\alpha_i\}_{i=1}^M} \sum_i \|\alpha_i\|_0 \text{ s.t. } \|y - D\alpha_i\|_2 \leq \varepsilon \quad (3)$$

The above minimization problem can be solved using K-SVD [13].

IV PROPOSED FUSION SCHEME

The schematic diagram of the proposed fusion framework is shown in Fig. 1. First, MST is applied to the source images $\{I_A, I_B\}$ to obtain their low-pass bands $\{L_A, L_B\}$ and high-passbands $\{H_A, H_B\}$. Max absolute rule[10] is used as a saliency measure to select the focused coefficients in the approximation bands. High frequency sub bands reflect the structural information of source images in horizontal, vertical and diagonal directions. Hence, SR based fusion rule is applied to obtain the fused high frequency bands.

A. High Pass Fusion

Merge H_A and H_B to obtain H_F with the popular "max-absolute" rule using the absolute value of each coefficient.

$$H_F = \begin{cases} H_A & \text{if } Abs_{HA}^i > Abs_{HB}^i \\ H_B & \text{otherwise} \end{cases} \quad (4)$$

B. Low Pass Fusion

1. Divide the low frequency bands L_A and L_B into image patches of size 8×8 from upper left to lower right with a step length of one pixel using sliding window technique. Suppose that there are T patches denotes in $\{P_{LA}^i\}$ and $\{P_{LB}^i\}$ where $1 < i < T$ L_A and L_B respectively

2. Rearrange the image patches $\{P_{LA}^i, P_{LB}^i\}$ into column vectors $\{V_{HA}^i, V_{HB}^i\}$

3. Calculate the sparse coefficient vectors $\{\alpha_{LA}^i, \alpha_{LB}^i\}$ of $\{V_{LA}^i, V_{LB}^i\}$ using OMP.

$$\alpha_{LA}^i = \arg \min \sum_i \|\alpha\|_0 \text{ s.t. } \|V_{LA}^i - D\alpha_i\|_2 \leq \varepsilon \quad (5)$$

$$\alpha_{LB}^i = \arg \min \sum_i \|\alpha\|_0 \text{ s.t. } \|V_{LB}^i - D\alpha_i\|_2 \leq \varepsilon \quad (6)$$

4. Merge α_{LA}^i and α_{LB}^i with the "max-L1" rule to obtain the fused sparse vector. It is given by,

$$\alpha_{LF}^i = \begin{cases} \alpha_{LA}^i & \text{if } \|\alpha_{LA}^i\| > \|\alpha_{LB}^i\| \\ \alpha_{LB}^i & \text{otherwise} \end{cases} \quad (7)$$

The fused vector is calculated by,

$$V_{LF}^i = D\alpha_{LF}^i \quad (8)$$

5. Iterate the above process for all the image patches to obtain

all the fused vector $\{V_{LF}^i\}_{i=1}^T$ For each V_{LF}^i reshape it into a patch and then merge all the patches to get the final fused high

pass coefficients. Perform inverse MST over the fused sub bands to reconstruct the final fused image.

III. EVALUATION METRICS

Image quality evaluation is done to know the degradation of a fused image with the reference image. There are different metrics used in evaluating the performance of the fusion methods like PSNR, SNR, SSIM, MAE. Abbreviations and Acronyms

1. Peak signal to noise ratio (PSNR):

PSNR determines the resemblance between reference image and fused image. Higher value of PSNR indicates good fusion results.

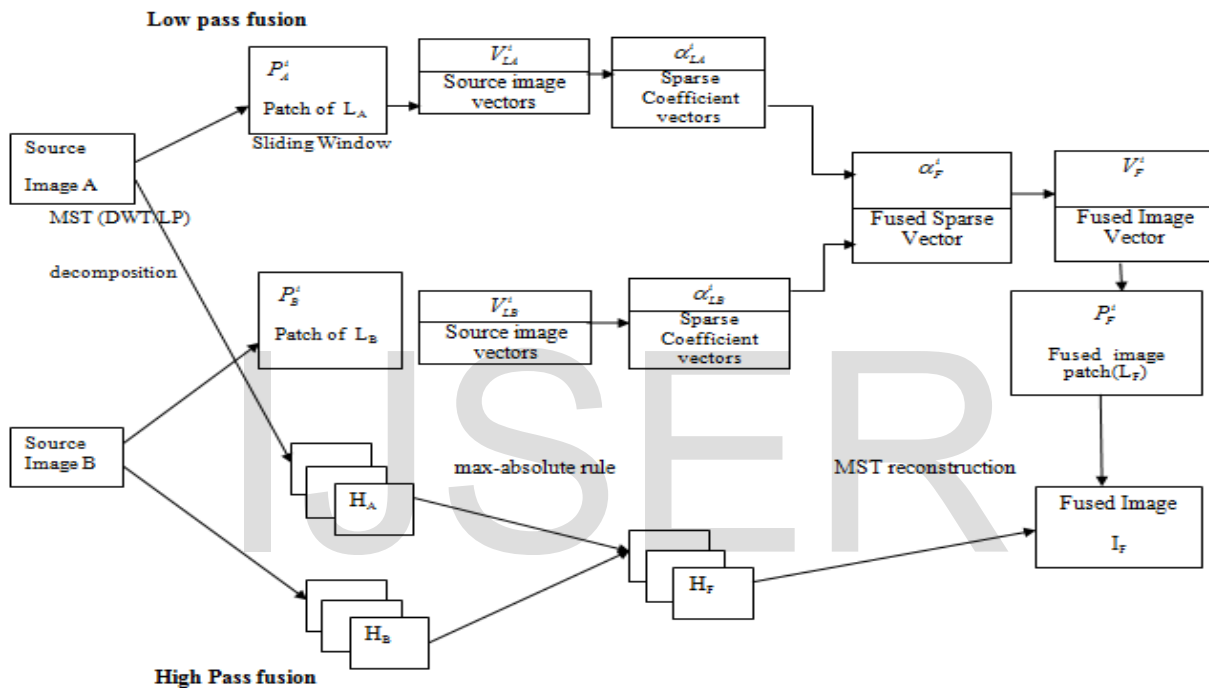


Fig 3: The schematic diagram of the proposed fusion framework.

$$\text{PSNR} = 20 \log_{10} \left[\frac{L^2}{\frac{1}{p \cdot q} \sum_{i=1}^p \sum_{j=1}^q (R(i, j) - F(i, j))^2} \right] \quad (9)$$

Here, L denotes number of gray levels in the image.

$R(i, j)$ is the reference image and $F(i, j)$ is the fused image.

2. Signal to noise ratio (SNR):

SNR gives the signal to noise ratio by comparing fused image with the reference image. Larger value of SNR indicates good fusion results.

$$\text{SNR} = 20 \log_{10} \left[\frac{\sum_{i=1}^p \sum_{j=1}^q R(i, j)}{\sum_{i=1}^p \sum_{j=1}^q (R(i, j) - F(i, j))^2} \right] \quad (10)$$

where $R(i,j)$ and $F(i,j)$ are reference and fused images respectively.

3. Mean Absolute Error(MAE):

MAE gives the mean absolute error between fused image and reference image. It is given by:

$$MAE = \frac{1}{p * q} \sum_{i=1}^p \sum_{j=1}^q |R(i, j) - F(i, j)| \quad (11)$$

4. Structural Similarity Index Metric(SSIM):

SSIM provides measure of structural information change of fused image from the reference image. SSIM index has a decimal value between 0 and 1. A value of 0 means zero similarity with original image and 1 means exactly same image as the original one. SSIM is given by:

$$SSIM = \frac{(2\mu_R\mu_F + c_1)(2\sigma_{RF} + c_2)}{(\mu_R^2 + \mu_F^2 + c_1)(\sigma_R^2 + \sigma_F^2 + c_2)} \quad (12)$$

where, μ_R and μ_F are mean intensities of images R and F.
 σ_R and σ_F are variances of images R and F.
 c_1 and c_2 are small constants of R and F respectively.
 $c_1 = (K_1, L)^2$ and $c_2 = (K_2, L)^2$, L is the dynamic range.
 $K_1 = 0.01$ and $K_2 = 0.03$ by default.

IV. RESULTS AND DISCUSSION

The performance of both the fusion methods are tested and compared. The test image used here is satellite image which is left focused and right focused by applying averaging mask are shown in 4(a & b). These multi focused images fused by the method shown above for levels ranging from one to four and the resultant image for each level is shown below. Fig 5(a-d) shows the fused images obtained after applying DWT and SR based method. Fig 6(a-d) shows the fused images obtained after applying LP and SR based method. However, it is hard to tell the difference between the results of the proposed method by subjective evaluation. Hence, in order to better evaluate these fusion methods, qualitative assessments of the performance of the two methods are needed. Four evaluation criteria including Peak Signal to Noise

Ratio(PSNR), Mean Absolute Error(MAE), Signal to Noise Ratio(SNR), Structural similarity Index(SSIM) are made use for quality assessment of fused images.

The above four evaluation criteria are then applied to evaluate the two fusion methods in Fig.5, and the detailed qualitative results for each level are given in Table 1. and Table 2. From Table 1, and Table 2 we can observe that the values of PSNR, SNR and SSIM is larger where as value of MAE is smaller of DWT and SR based method compared to LP and SR based method, which means that DWT and SR based method performs better than the LP and SR based method.

V. CONCLUSION

In this paper, we proposed a novel multifocus image fusion algorithm using MST and sparse representation. Initially, the low frequency components and high frequency components are separated using MST. Max absolute rule is used as a focus measure to fuse the high frequency sub bands. SR based fusion rule is applied to obtain the fused low frequency coefficients.

Finally, inverse MST is applied to get the fused image. In our experiment, two popular multi-scale transforms (LP, DWT) with different decomposition levels ranging from one to four are employed for the fusion of multi-focus images.

Also we have incorporated parameters like peak signal to noise ratio(PSNR), mean absolute error(MAE), signal to noise ratio (SNR) and structural similarity index(SSIM) to evaluate both the methods. As shown in the above result, the values of PSNR, SSIM, SNR is larger where as MAE value is less for DWT and SR based method when compared to LP and SR based method which proves that DWT performs better than LP.



Fig 4: a) Right focused image b) Left focused image

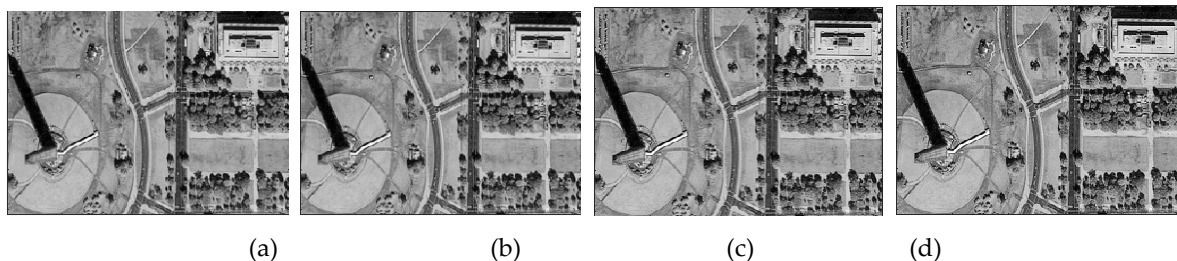


Fig 5: Fusion result using DWT and SR based method for different levels of decomposition a) Level 1 b) Level 2 c) Level 3 d) Level 4

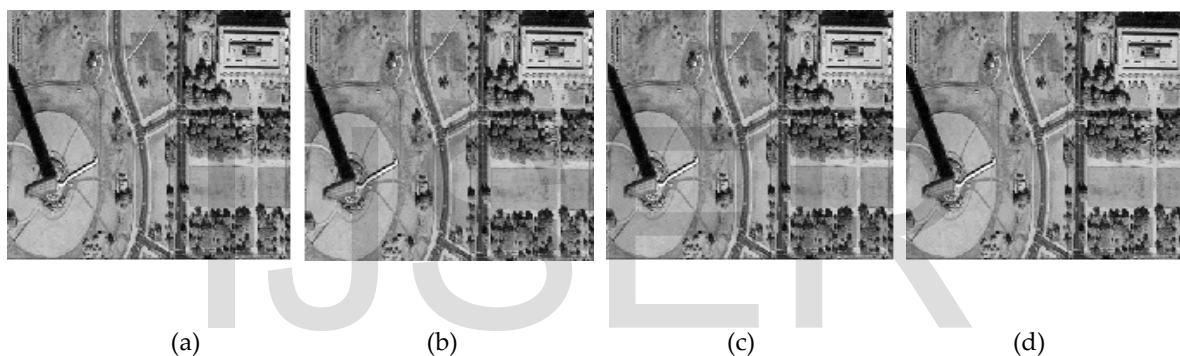


Fig 6: Fusion result using LP and SR based method for different levels of decomposition a) Level 1 b) Level 2 c) Level 3 d) Level 4

Parameters /Levels	Level 1	Level 2	Level 3	Level 4
PSNR	42.6457	38.07	36.9577	36.5551
MAE	0.99769	0.99387	0.99291	0.99298
SNR	37.5938	33.0181	31.9058	31.5032
SSIM	0.1292	0.32791	0.52312	0.61632

Table 1: Qualitative assessment results for DWT and SR based method

Parameters /Levels	Level 1	Level 2	Level 3	Level 4
PSNR	40.7079	39.5643	38.9052	38.6079
MAE	0.16553	0.21378	0.39764	0.41789
SNR	35.656	34.5123	33.8532	33.556
SSIM	0.99547	0.99458	0.994446	0.99442

Table 2: Qualitative assessment results for LP and SR

REFERENCES

- [1] Yong Yang. A Novel DWT Based Multi-focus Image Fusion Method. *International Conference on Advances in Engineering* 2011; 24: 177-181.
- [2] Nirmala Paramanandham, Kishore Rajendiran, Deepika Narayanan, Indu Vadhani.S, Mrinalini Anand. An efficient Multi transform based Fusion for Multi Focus Images. *IEEE ICCSP conference*. 2015.
- [3] G.Madhuri, Dr.CH.Hima Bindu. Performance Evaluation of Multi-Focus Image Fusion Techniques. *Intl. Conference on Computing and Network Communications (CoCoNet'15)*, 2015.
- [4] Miss. Devyani P. Deshmukh, Prof. A. V. Malviya. Image Fusion an Application of Digital Image Processing using Wavelet Transform. *International Journal of Scientific & Engineering Research; Volume 6, Issue 11, November-2015*
- [5] Wencheng Wang A Multi-focus Image Fusion Method Based on Laplacian Pyramid. *Journal of computers*, vol. 6, no. 12, december 2011
- [6] Dr. H.b. kekre, Dr. Dharendra mishra, Rakhee saboo. Review on image fusion Techniques and performance Evaluation parameters. *International Journal of Engineering Science and Technology (IJEST)*; Vol. 5 No.04 April 2013
- [7] Shriniwas T. Budhewar. Wavelet and Curvelet Transform based Image Fusion Algorithm. *(IJCSIT) International Journal of Computer Science and Information Technologies*, Vol. 5 (3) , 2014
- [8] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. Huang, S. Yan, "Sparse representation for computer vision and pattern recognition", *Proc. IEEE*, vol. 98, pp. 1031–1044, 2010.
- [9] M. Elad, M. Figueiredo, Y. Ma, "On the role of sparse and redundant representations in image processing", *Proc. IEEE*, vol. 98, pp. 972– 982, 2010.
- [10] H. Li, B. Manjunath, S. Mitra, Multisensor image fusion using the wavelet transform, *Graphical Models and Image Processing* 57 (3) (1995) 235–245.
- [11] B. Yang, S. Li, "Multifocus image fusion and restoration with sparse representation", *IEEE Trans. Instrum. Meas.*, vol. 59, pp. 884–892, 2010.
- [12] Chunling Zhao , Quanxin Ding and Jianxun Li "The performance analysis of image fusion algorithm *International Symposium on Computational Intelligence and Design* 2008.
- [13] M. Elad, M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries", *IEEE Trans. Image Process.*, vol. 15, pp. 3736-3745, 2006.
- [14] M. Aharon, M. Elad, A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation", *IEEE Trans. Signal Process.*, vol. 54, pp. 4311–4322, 2006.