Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 13, No. 1, (2022),
ISSN : **1906-9685**

MEDICAL IMAGE FUSION BASED ON MULTISCALE TRANSFORM WITH SPARSE REPRESENTATION

[1]Prathap Kumar M, [2] Syam Prasad Pamula, [3] R.Anji Naik, [4]Kuriti Chandu
[1,2,3]Assistant Professor, [4]Student, Dept. of Electronics & Communication Engineering, Newton's Institute of Engineering, Macherla, Andhra Pradesh, India.

ABSTRACT

There are two commonly used techniques in image fusion techniques: multi-scale transform (MST) and sparse representation (SR). This paper presents the combination of MST and SR to overcome the defects in the MSF-GIF-based fusion technique. In our method, the MST is firstly performed on both MRI images and PET images and then we obtain low-pass and high-pass coefficients. The low-pass coefficients are dealt with SR base fusion methods. We have done the MST decomposition using discrete wavelet transform. Finally, we get our final fused image by doing an inverse multi scale transform. The advantages of our proposed method are it exhibits detailed information compared to the experimental results shown by MSF-GIF final results. By comparing the parameters of fused images subjectively and objectively, we get the best-performed fusion method under the proposed framework for each category of image fusion. Furthermore, experimental results show that the proposed fusion method can obtain good performance for the fusion of multimodal images.

Keywords: Image, MST, Sparse representation

1. INTRODUCTION

Image fusion has become a contentious issue in the image processing field in recent years. The goal of image fusion is to create a composite image by combining complementary information from numerous source images of the same scene into a single composite image. It is challenging to obtain an image that contains all of the image's important information. Along with the required image, an image may contain background images.

Due to its restricted focusing skills, the background and needed content may not be caught. A very significant element is capturing an image with all of the information from the original image. There are numerous image fusion methods that have a wide range of applications in medical imaging, satellite imaging, remote sensing, and other fields. Medical imaging applications are taken into account in our project.

Image fusion is a technique for extracting information from many images and combining it into a single image to improve perception. Fused images are more informative and more compatible with human sight. As a result, they're used in a variety of applications to improve their performance, including medical imaging, surveillance, military, and remote sensing. Multi-focus, multi-view, and multi-modal situations are examples of fusing scenarios.

http://doi.org/10.36893/JNAO.2022.V13N1.841-848

2 PRELIMINARIES

PRELIMINARY CONCEPTS

2.1 IMAGEFUSION

Image fusion is the process of combining many photos from various sources with their corresponding complementary information to create a new image that contains all of the common and complementary elements of the original images. Multisensory systems have become a reality in a variety of sectors, including remote sensing, medical imaging, machine vision, and military applications, thanks to recent rapid advances in imaging technologies. By extracting all of the useful information from the source photos, image fusion is an effective technique for decreasing the growing number of data. Image fusion is a useful tool for comparing and analyzing multi-sensor data with complementary information about the area in question. Image fusion generates new images that are better suited to human and machine perception.The goal of picture fusion is to create images that are more appropriate and intelligible for human and machine perception, not just to minimize the amount of data.

Multisensor image fusion is a computer vision technique that combines pertinent information from two or more images into a single image.We combined medical photos in this project.

AnAccurate diagnosis requires data from many modality images, such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Positron Emission Tomography (PET), and others, which can be gathered using the Image Fusion method.

Fig.1. Target Detection using Image Fusion

2.2MULTISCALE TRANSFORM

For the fusion of MST low-pass bands, the sparse coding approach is used in this study. Sampling from the matching MST low-pass bands acquired from some training images under the same decomposition condition is one technique to get the training patches.

We hope to learn a universal vocabulary that can be used in any transform domain and parameter settings in this project. The MST low-pass band created by image filtering can be thought of as a smooth version of the original image, as is widely known. Because a dictionary learned from natural image patches can sparsely represent the various "flat" patches found in a natural image.

In order for an input patch to be represented, it must also have a mean value.

2.3 IMAGE DECOMPOSITION

The Wavelet Transform converts a signal into a time-frequency representation. It was created to address a flaw in the Short Time Fourier Transform (STFT), which can be used to examine non-stationary signals as well. While the STFT maintains a consistent resolution across all frequencies, the Wavelet Transform employs a multi-resolution approach that analyses different frequencies with different resolutions.

The Discrete Fourier Transform (DFT) decomposes a signal into sinusoidal basis functions of various frequencies in Fourier analysis. This transformation does not lose any information; in other words, the original signal may be fully recovered from its DFT (FFT) representation. The Discrete Wavelet Transform (DWT) decomposes a signal into a set of mutually orthogonal wavelet basis functions in wavelet analysis.

These functions differ from sinusoidal basis functions in that they are spatially localised, meaning that they are nonzero only across a portion of the whole signal duration. Wavelet functions are also dilated, translated, and scaled variants of the mother wavelet, which is a common function.The DWT is invertible, just like Fourier analysis, therefore the original signal can be reconstructed completely from its DWT form.

Unlike the DFT, the DWT refers to a collection of transforms, each having its own set of wavelet basis functions. The Haar wavelets and the Daubechies wavelets are two of the most popular. Figures 1 and 2, for example, show the complete set of 64 Haar and Daubechies-4 wavelet functions. We will not get into the intricacies of how these were generated here; nonetheless, the following crucial qualities should be noted:

1. Wavelet functions are spatially localized; 2. Wavelet functions are dilated; 3. Wavelet functions are dilated; 4. Wavelet functions are dilated; 5. Wavelet functions are dilated

2.4SPARSE REPRESENTATION:

Sparse representation theory proposes a new, extremely effective, and ubiquitous paradigm of this type. Its key concept is that data is described as a linear combination of a few building pieces - atoms - selected from a pre-defined dictionary of such fundamental elements.

3 PROPOSED METHODOLOGY BLOCK DIAGRAM

The Image Fusion process involves some methods to obtain the final result and they are mentioned in the block diagram.

Different images with different image principles are used which are given as source images. The images used were x-ray, MRI images, and PET images.

For the fusion of MST low-pass bands, the sparse coding approach is used in this study. Sampling from the matching MST low-pass bands acquired from some training images under the same decomposition condition is one technique to get the training patches.

We hope to learn a universal vocabulary that can be used in any transform domain and parameter settings in this project. The MST low-pass band created by image filtering can be thought of as a smooth version of the original image, as is widely known. Because a dictionary learned from natural image patches can sparsely represent the various "flat" patches found in a natural image.

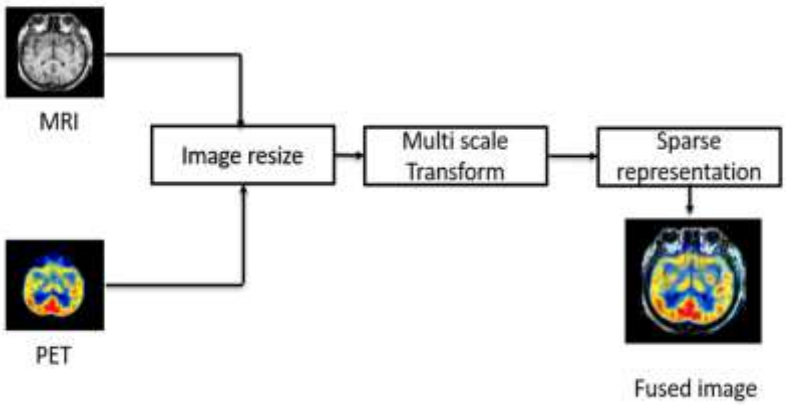


Fig.2. Block Diagram showing image fusion process

3.1 IMAGE RESIZE

When you enlarge or distort an image from one-pixel grid to another, you're using image interpolation. When you need to increase or decrease the total amount of pixels in an image, you'll need to resize it, whereas you'll need to remap it if you're correcting for lens distortion or rotating it.

3.2 Multi scale transform

Step1: MST decomposition- perform a specific MST on the two source image $\{I_A, I_B\}$ to obtain their low pass bands $\{L_A, L_B\}$ and high pass bands which are uniformly denoted as $\{H_A, H_B\}$.

Step2: Low pass image fusion

i. apply sliding window technique to divide L_A and L_B into image patches of size $\sqrt{n} \times \sqrt{n}$ from upper left to lower right with a step length of s pixels.
ii. For each position i rearrange $\{p_A^i, p_B^i\}$ into two column vectors $\{v_A^i, v_B^i\}$ and then normalize each vectors mean value to zero.
iii. Calculate the sparse coefficient vectors
iv. Merge α_A^i and α_B^i with the ''max-L1'' rule to obtain the fused sparse vector.
v. Iterate the above process for all the source image patches to obtain all the fused vectors.
Step 3: 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 as the activity level measurement. Then, apply the consistency verification scheme to ensure that a fused coefficient does not originate from a different source image from most of its neighbors. This can be implemented via a small majority filter.
Step 4: MST reconstruction.
Perform the corresponding inverse MST over LF and HF to reconstruct the final fused imageI_F.

3.3 SPARSE REPRESENTATION

With the help of the complete functions accessible in MATLAB, the SR-based methodology calculates the coefficients of a sparse matrix. Then, using the specified max absolute methodology, these coefficients are blended. Finally, the merged image is reconstructed using the functions from the merged coefficients of the SR representation.
Using an over-complete dictionary D, the signal x can be represented as a linear combination of dictionary atoms in the sparse representation theory
x=Dx=D (1)

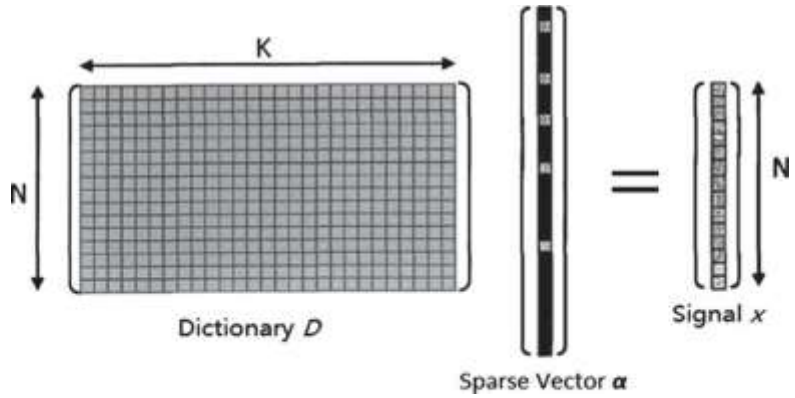


Fig.3. sparse representation matrix

Sparse representations are created using a basis that takes advantage of the input signals' regularity, resulting in a large number of small-amplitude coefficients. Because wavelets have localized support, functions with isolated singularities create a small number of large-amplitude wavelet coefficients in their vicinity. Over spaces of functions with "not too many" sharp transitions and singularities, nonlinear wavelet approximation gives a minimalinaccuracy. Functions with a restricted total variation norm are suitable models for images with nonfractal (finite length) edges.

4. Advantages of the proposed method over individual MST or SR

The MST-based fusion method has a number of disadvantages, including low-band averaging and the loss of the majority of the image energy. referred to in the first part,An illustration will help to clarify the situation. The multi-modal pictures seen in are made up of a pair ofpictures visible and MMW The multi-modal pictures' source images Varied sensors contribute to the fusion, and each sensor receives a different amount of data. from the same scene, various physical traits As much as is possible. The visible image, as seen in Fig. 4, reveals edges and textures, whereas the hidden image does not. A bright patch in the image can be used to detect the hidden metallic object.the millimeter image. The sum of grey levels squared is a number. A measurement of the image's energy the higher the grey levels, the better. The brighter regions were the result of the energy.

By using SR on the low-pass bands, the proposed strategy can fix the problem. In fact, in the SR approach, subtracting the mean value from the patches maintains the majority of the source image's energy and transfers it to the final fused image. Determining the decomposition levels of the MST-based fusion approach is another challenge. A significant number of breakdown stages must be chosen in order to retrieve enough features. Low-band coefficients, on the other hand, can occur, affecting a large number of pixels. As a result of the misregistration, a small noise or error appears in the source image, resulting in a distorted fused image. In contrast, the suggested method uses the SR method to extract low band spatial information, and four decomposition levels are sufficient.

The following is how our proposed technique overcomes the problems mentioned:

i. The SR approach is applied to the image's low band. Because there are no critical features in the low band and they are truly in the upper bands, sliding widow might choose large step. As a result, the 'blocking effect' becomes less effective.

ii. The low band of a picture in most MSTs is much smaller than the source image. As a result, the cost of computing SR is greatly decreased, and the second problem is eliminated.

5. EXPERIMENTAL RESULTS

5.1 OBJECTIVE IMAGE QUALITY METRICS:

The objective evaluation approach calculates and obtains numerical values for the fused images using mathematical procedures, which measure the degree of distortion between the fused and input images. For measuring the retention of edges and textures, the amount of detail contrast, aIt is used to measure the overall activity level of the space in the image. The larger the **SF**of the fused image, the richer the information it contains, and the better the fusion method has performed. It is defined as follows:

$$SF = \sqrt{RF^2 + CF^2} \quad (2)$$

Natural image quality evaluator (NIQE):it evaluates the quality of the fused images when the type of fused image distortion is unknown. The smaller the NIQE,the higher the fusion performance. It is defined as follows:

$$\text{NIQE} = \sqrt{(\mu1 - \mu2)^T . (\frac{\sigma1+\sigma2}{2})^{-1} . (\mu1 - \mu2)} \quad (3)$$

Tone-mapped image quality Index (TMQI): It is an index for measuring the degree of loss of the fused image contrast information and luminance information. The larger the value is, the better the fusion result.

$$TMQI(I_R, I_F) = aT^{\alpha} + (1 - a)M^{\beta} \quad (4)$$

where T and M represent the structural fidelity and the statistical properties, respectively, of the image and the values of the constants are a=0.8012, α =0.3046, and β =0.7088.

Edge intensity (*EI*):It is used to detect the edges of the image and a larger value of *EI* corresponds to richer edge information in the image.

$$\mathbf{EI} = \frac{\mathbf{1}}{\boldsymbol{M \times N}} \times \sum_{\boldsymbol{j=1}}^{\boldsymbol{N}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{M}} \sqrt{\boldsymbol{G_x(i,j)^2 + G_y(i,j)^2}} \quad (5)$$

Mutual information (MI): It is used to calculate the information correlation between the fused image and the input source images. The larger the value of *MI*, the larger the amount of source image information that is contained in the fused image.

$$\mathbf{MI}(\boldsymbol{I_R}, \boldsymbol{I_F}) = \frac{\boldsymbol{H(I_R)+H(I_F)-H(I_R,I_F)}}{\boldsymbol{H(I_R)+H(I_F)}} \quad (6)$$

Edge-dependent fusion quality index (*Qe*):It is used to measure the edge information.

$$Q_e(I_R, I_F) = w . Q_{wb}(I_1, I_F) + (1 - w) \times Q_{wb}(I_2, I_F) \quad (7)$$

where Q_{wb}denotes average value between the reference image and the fused image. The larger the value of Q_e, the better the fused image.

Entropy (*EN*):It measures the information in the fused image.

$\mathbf{EN}(I_F) = -\sum_{x=0}^{255} p_x(I_F) \ln p_x(I_F)$ (8)

$Q_{ab/f}$: It measures the success of edge information transfer from the input images to the fused image.

Quality-aware clustering (*QAC*)

It is used to predict features via machine learning. The smaller the *QAC* value is, the less distortion in the fused image. *QAC* consists of two stages: training and testing. In the stage of training, distortion models, such as Gaussian noise, blur, JPEG, and JPEG 2000 compression, are trained using machine learning methods. In the stage of testing, *QAC* is calculated as the sum of *Zl*:

$\mathbf{QAC}(I_F) = \frac{1}{L}\sum_{l=1}^{L} z_l$ (9)

j. Average gradient (*AG*)

It represents the details of the image contrast and texture features using horizontal gradient and vertical gradient

$\mathbf{AG}(I_F) = \frac{1}{(M-1)(N-1)} \times \sum_{i=1}^{M-1}\sum_{j=1}^{N-1}\sqrt{\frac{(\Delta_x I_F(i,j))^2 + (\Delta_y I_F(i,j))^2}{2}}$ (10)

k. Spatial-spectral entropy-based quality (*SSEQ*)

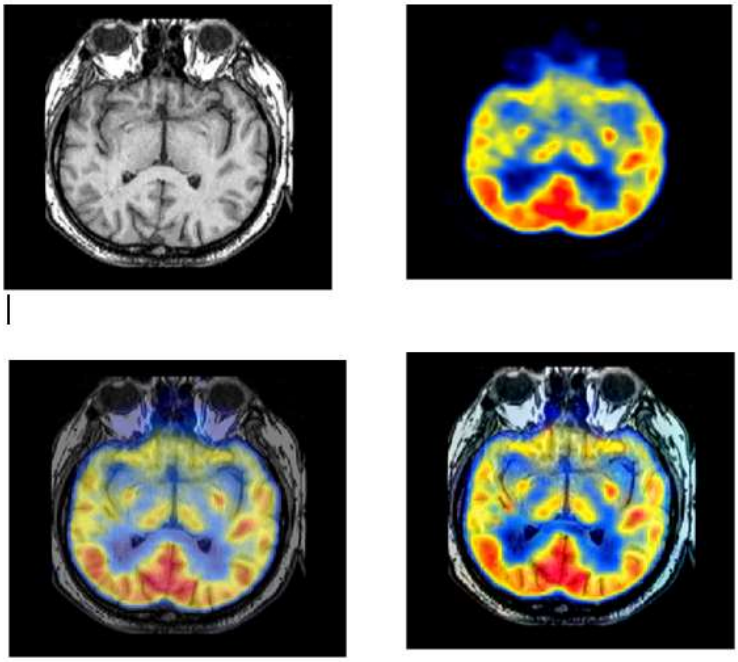
It is used to evaluate the quality of a distorted image across unknown distortion groups. *SSEQ* utilizes local spatial entropies and local spectral entropies to predict the image quality score. The distorted image is predicted using a 12-dimensional feature vector *f*.

SSEQ is calculatedby applying the function libsvmon the feature vector $\boldsymbol{SSEQ = libsvm(f)}$ (11)

5.2 OBJECTIVE EVALUATION:

PARAMETERS	MSF-GIF	MST-SR
SF	0.094442	**0.097174**
NIQE	61.482414	**58.114560**
TMQI	0.650795	**0.965350**
EI	89.177473	**93.203223**
MI	5.292621	**6.897078**
QE	0.993626	**0.994733**
EN	5.039720	**5.437674**
QAC	0.514461	**0.415559**
AG	0.039708	**0.049143**
Qab/f	0.731523	**0.894119**
SSEQ	0.333208	**0.328313**

5.3 SUBJECTIVE EVALUATION



Input MRI ImageInput PET Fused MSF-GIF IMAGE Fused MST- SR Image

Fig. 4: Subjective experimental results

CONCLUSION

The MST-SR fusion approach is used in this paper to present an effective multimodal medical picture fusion solution. The results show that fusion based on MST-SR produces better results than fusion based on guided filter and salient features extraction. Because the fusion products have more complete structures and colour features, as well as crisper edges and textures, the experimental results show that the method can be effectively applied to MRI-PET fusion.

REFERENCE

[1] J. Du, W. Li, K. Lu, and B. Xiao, ``An overview of multi-modal medicalimage fusion," *Neurocomputing*, vol. 215, pp. 3_20, Nov. 2016.

[2] C. He, Q. Liu, H. Li, and H. Wang, ``Multimodal medical image fusion based on IHS and PCA," *Proc. Eng.*, vol. 7, pp. 280_285, Jan. 2010.

[3] S. Daneshvar and H. Ghassemian, ``MRI and PET image fusion by combining IHS and retina-inspired models," *Inf. Fusion*, vol. 11, no. 2, pp. 114_123, 2010.

[4] M. J. Li, Y. B. Dong, and X. L. Wang, ``Image fusion algorithm based on gradient pyramid and its performance evaluation," *Appl. Mech. Mater.*, vol. 525, pp. 715_718, Feb. 2014.

[5] P. J. Burt and E. H. Adelson, ``The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532_540, Apr. 1983.

[6] V. Bhateja, H. Patel, A. Krishn, A. Sahu, and A. Lay-Ekuakille, ``Multimodal medical image sensor fusion framework using cascade of wavelet and contourlet transform domains," *IEEE Sensors J.*, vol. 15, no. 12, pp. 6783_6790, Dec. 2015.

[7] G. Pajares and J. M. de la Cruz, ``A wavelet-based image fusion tutorial," *Pattern Recognit.*, vol. 37, no. 9, pp. 1855_1872, 2004.

[8] B. Yang and S. Li, ``Pixel-level image fusion with simultaneous orthogonal matching pursuit," *Inf. Fusion*, vol. 13, no. 1, pp. 10_19, 2012.

[9] Y. Liu, S. Liu, and Z.Wang, ``A general framework for image fusion based on multi-scale transform and sparse representation," *Inf. Fusion*, vol. 24, pp. 147_164, Jul. 2015.

[10] H. Cheng, Z. Liu, L. Yang, and X. Chen, ``Sparse representation andlearning in visual recognition: Theory and applications," *Signal Process.*, vol. 93, no. 6, pp. 1408_1425, 2013.

[11] S. Li, H. Yin, and L. Fang, ``Group-sparse representation with dictionary learning for medical image denoising and fusion,'' *IEEE Trans. Biomed.Eng.*, vol. 59, no. 12, pp. 3450_3459, Dec. 2012.

[12] Y. Liu, X. Chen, J. Cheng, and H. Peng, ``A medical image fusion method based on convolutional neural networks,'' in *Proc. 20th Int. Conf. Inf.Fusion*, Jul. 2017, pp. 1_7.

[13] S. Monica and S. Sahoo, ``Pulse coupled neural networks and its applications,'' *Expert Syst. Appl.*, vol. 41, no. 8, pp. 3965_3974, 2014.

[14] X. Hou and L. Zhang, ``Saliency detection: A spectral residual approach,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1_8.

[15] N. Imamoglu, W. Lin, and Y. Fang, ``A saliency detection model using low-level features based on wavelet transform,'' *IEEE Trans. Multimedia*, vol. 15, no. 1, pp. 96_105, Jan. 2013.

[16] M. Song, C. Chen, S.Wang, and Y. Yang, ``Low-level and high-level prior learning for visual saliency estimation,'' *Inf. Sci.*, vol. 281, pp. 573_585, Oct. 2014.

[17] J. Harel, C. Koch, and P. Perona, ``Graph-based visual saliency,'' in *Proc. 19th Int. Conf. Neural Inf. Process. Syst.*, 2007, pp. 545_552.

[18] L. Zhang, A. Li, Z. Zhang, and K. Yang, ``Global and local saliency analysis for the extraction of residential areas in high-spatial-resolution remote sensing image,'' *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 3750_3763, Jul. 2016.

[19] C.-L. Chien and W.-H. Tsai, ``Image fusion with no gamut problem by improved nonlinear IHS transforms for remote sensing,'' *IEEE Trans.Geosci. Remote Sens.*, vol. 52, no. 1, pp. 651_663, Jan. 2014.

[20] S. Chen, R. Zhang, H. Su, J. Tian, and J. Xia, ``SAR and Multispectral image fusion using generalized IHS transform based on à Trous wavelet and EMD decompositions,'' *IEEE Sensors J.*, vol. 10, no. 3, pp. 737_745, Mar. 2010.

[21] G. Bhatnagar, Q. M. J. Wu, and Z. Liu, ``Human visual system inspired multi-modal medical image fusion framework,'' *Expert Syst. Appl.*, vol. 40, pp. 1708_1720, Apr. 2013.

[22] S. Li, X. Kang, L. Fang, J. Hu, and H. Yin, ``Pixel-level image fusion: A survey of the state of the art,'' *Inf. Fusion*, vol. 33, pp. 100_112, Jun. 2017.

[23] P. Gomathi and B. Kalaavathi, ``Medical image fusion based on multiscale transforms,'' *J. Med. Imag. Health Inform.*, vol. 7, no. 2, pp. 478_484, 2017.

http://doi.org/10.36893/JNAO.2022.V13N1.841-848