1st International Conference on Applied Soft Computing Techniques
22 & 23.04.2017
In association with
International Journal of Scientific Research in Science and Technology

Multi-focus Image Fusion with Quantitative Analysis

S. Abirami, G. Rajasekaran
Department of Information Technology, Mepco Schlenk Engineering College, Sivakasi, TamilNadu , India

ABSTRACT

Multi-focus image fusion is the combining relevant information from two or more images of a same scene and as results has "all-in-focus" image. When one scene contains objects in different distance, the camera can be focused on one another of each object, generate set of pictures. Then, by applying image fusion techniques, an image with better focus across all area can be generated. This paper describes an image fusion system using different fusion techniques and the resultant is analyzed with quantitative measures. Initially, the registered images from two different modalities are considered as input image. For the resultant data the perceptual image fusion is applied and the fused image is analyzed with quantitative metrics namely Peak Signal –to- Noise Ratio (PSNR), Mutual Information (MI), Structural Similarity Index (SSIM). From this experimental result we observed that the proposed fusion method provides better result compared to the given images as justified by quantitative metrics.

Keywords: Image Fusion; Discrete Wavelet Transform; Dual Tree Complex Wavelet Transform; Quantitative Metrics

I. INTRODUCTION

The effective fusion for more than one resource with high resolution and having high information content in context with visualization, scene understanding, target recognition and situational awareness in multi-sensor applications such as medicine, surveillance and remote sensing. In the research of fusion techniques, there is several fusion techniques are proposed and implemented, but these techniques have certain limitations. Some of the image fusion techniques are available in Fusion Tool of Matlab5.0 namely are Filter-subtraction-Decimate Pyramid (FSD), Gradient Pyramid, Laplacian Pyramid [4], Discrete Wavelet Transform Pyramid (DWT), Shift Invariant Discrete Wavelet Transform Pyramid (SIDWT) [3], Principle Component Analysis [2], Morphological Pyramid, Contrast Pyramid, Ratio Pyramid, and so on [1]. For example, Contrast Pyramid loses too much information from the source images; Lots of false information produced by the Ratio Pyramid method which never existed in the original images; Many false edges created by the Morphological Pyramid method. In a word, these methods cannot deal with various types of images. Many of the fusion techniques are based on wavelet transformation. But, the DWT image fusion method is resulting with shift variant and additive noise in fused image. Using Redundancy Discrete Wavelet Transform (RDWT), Contourlet Transform [5] and Dual-Tree Complex Wavelet Transform (DTCWT) [6]. An RDWT fusion method is used to preserve the exact edge and spectral information from the given images without any loss of spatial information. In this technique, the high pass and low pass sub bands of the input images are fused using the average method and entropy method respectively. The region based Contourlet Transform gives local brightness, localization, multiresolution, directionality and anisotropy, etc. on the fused image. This transformation process is implemented in two stages: a) transformation and b) subband decomposition. On the first stage, double filter bank is applied and the second stage, local energy is calculated to the each subband and then fusion rules are applied like average mode and selection mode. DTCWT has good directional selectivity as compare to other methods and it also reduced shift variant property.

Following are features of DTCWT:

- Good directional selectivity
- Approximate shift variant
- Perfect reconstruction using short linear filters
- Limited redundancy
- Efficient order n computations

The actual fusion process should retain as much perceptually important information as possible from the two sources and should form a single more informative image [1]. In the transform domain, firstly the input images are decomposed based on transform coefficients. Then the fusion technique is applied to obtain the fusion decision map. The inverse transformation on this decision map yields the fused image. The fused image contains all the information of the source images and reduces the spatial distortion.

The performances of different image fusion methods are analysed using different measures like quantitative metrics namely PSNR, SSIM and MI.

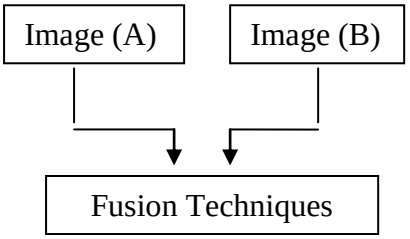
The rest of the parts are follows: In section II, the logical design is briefly reviewed; experimental results are described in section III; the performance of the proposed method is evaluated based on the quantitative metrics; section IV contains the conclusion.

II. LOGICAL DESIGN

In this system initially the two images are taken as input. Then the fusion techniques are applied to the registered images to find a more informative fused image. Finally, fused image information is validated using the quantitative measures. The overall system design is shown in Fig.1.

A. Dataset

The fusion image used by petrovic [11] is taken as the dataset (comprising remote sensing, medical and visual/thermal pairs). Each set of given images were taken from the same scene but at the different views.



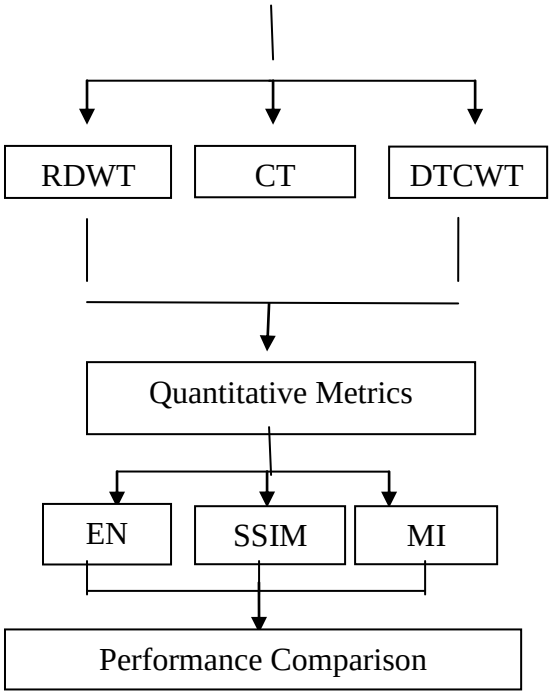


Figure 1. An overall view of logical design

B. Redundancy Discrete Wavelet Transform

RDWT fusion technique can be effectively performed on two registered images with captured at different time instance.

Let A and B be the two registered images of two different modalities. The registered images are decomposed into three levels of RDWT decomposition using Daubechies filters on the both input images in order to produce an approximate wavelet bands as shown in Fig. 2

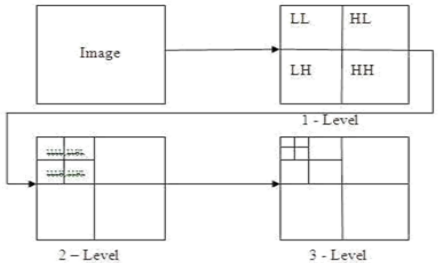


Figure 2. Three levels of decomposition

In the first phase I1 image is decomposed into $I_A^p, I_A^q, I_A^r, I_A^s$ be the RDWT subband and also $I_B^p, I_B^q, I_B^r, I_B^s$ be the corresponding RDWT subbands from I2 image. To extract the features from both the images, coefficients from approximation band of IA and IB are averaged;

$$I_N^a = mean(I_A^p, I_B^q) \quad (1)$$

Where I_N^a is the approximation band of the fused image.

In the next phase each subband namely LH, HL, HH, HL is divided into blocks of size 3*3 and the entropy of each block is calculated, as in (2)

$$e_i^{ab} = \ln \sqrt{(\mu_i^{ab} - \Sigma_{x,y=1}^{3,3} I_i^{ab}(x,y)/\sigma_i^{ab})^2 / r^2} \quad (2)$$

Where (a = q, r, s) specify the subbands, m = 3 (size of each block), b represents the block number, and (i = A, B) is used to differentiate the two multimodal images I_A and I_B . μ_i^{ab} and σ_i^{ab} are the mean and standard deviation of the RDWT coefficients. Using the entropy values, the detail subbands for the fused image I_N^q, I_N^r, I_N^s, and I_N^h are generated, the derived fused image block I_N^{ab}, RDWT coefficients from I_A image is greater than the specific block of I_B image, otherwise I_B^{ab} is selected.

$$I_N^{ab} = \begin{cases} I_A^{ab}, & (e_1^{ab} > e_2^{ab}) \\ I_B^{ab}, & otherwise \end{cases} \quad (3)$$

Finally, IRDWT is applied on all the subbands to generate the resultant fused image I_N.

$$I_N = IRDWT(I_N^p, I_N^q, I_N^r, I_N^s) \quad (4)$$

C. Contourlet Transform

The contourlet transform is a two-dimensional transform method for image representations. The CT has properties of multiresolution, directionality, localization, critical sampling and anisotropy. Its basic functions are multidimensional and multiscale.

Contourlet transform causes smoothness in a fused image with any two different modalities of images [6]. This transformation process involved in two stages. Double filter bank is applied for transformation in the first stage and decomposition process is done with fusion rules in the second stage. Finally, the fused output image is efficiently retrieved by using reconstruction procedure.

The block diagram of the proposed image fusion algorithm is shown in Fig. 3. Here images A and B represent the input source image. F is the final outcome of the fused image after applying the inverse contourlet transform.

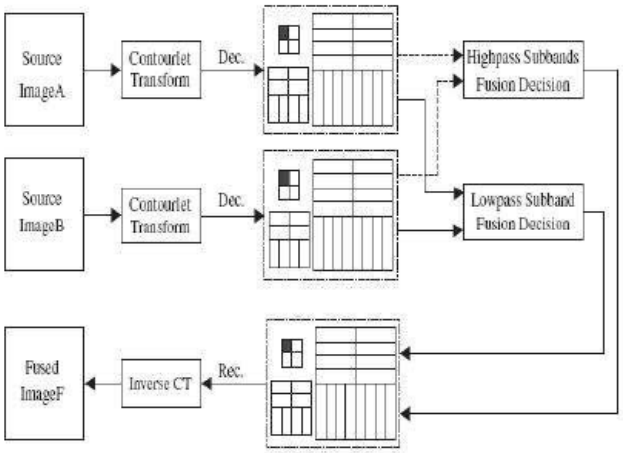


Figure 3. Flow diagram of the contourlet based image fusion

1) Transformation Stage

A double filter bank scheme is utilized efficiently for subband decomposition using Laplacian Pyramid (LP) and Directional Filter Bank (DFB) in this stage. To capture the edge point Laplacian Pyramid is used. To link the discontinuities point in linear structures Directional Filter Bank is used.

Each input image is decomposed into a subband of low frequency of original image and a bandpass high frequency subbands in Laplacian Pyramid method [7]. The same process is repeated for the specified contourlet decomposition level for low frequency subband. The decomposition process of Laplacian Pyramid is shown in Fig. 4. First the input image is applied to a LP filter H and then down sampled to derive a coarse approximation a. After performing the down sampled the image is up sampled and passed through a synthesis filter G. The result of highpass subbands are derived from subtracting the output of the synthesis filter with the input image.

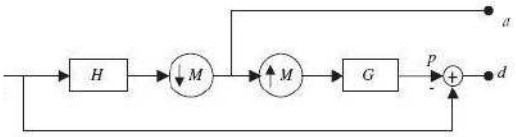


Figure 4. Construction of LP

2) Decomposition Stage

The decomposed subbands of transformation stage are merged using lowpass and highpass fusion rule.

a) Lowpass subband fusion:

The coefficients in the coarest scale subband a represents the approximation component of the source image. In this method, the local energy contourlet domain is developed as the measurement, then the selection mode and averaging modes are used to compute the final coefficients [7].

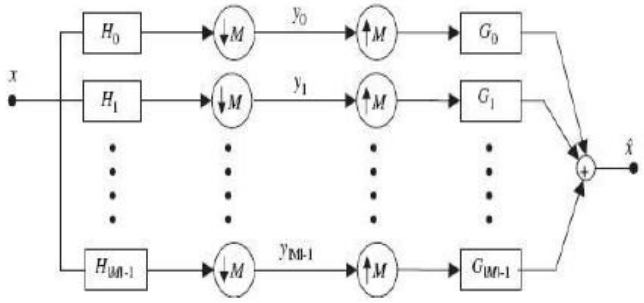


Figure 5. Construction of DFB

The local energy E(x,y) is calculated the current Coefficient in the approximate subband a, which is

$$E(x,y) = \Sigma_m \Sigma_m a_j(x+m, y+n)^2 W_L(m,n) \quad (5)$$

Where (x,y) is the current contourlet coefficient, $w_L(m,n)$ is a template of size 3*3

$$W_L = \frac{1}{9} * \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (6)$$

The salience factor is calculated to identify whether the selection mode and averaging mode to be used in the fusion process.

$$M_j^{AB}(x,y) = 2\sum_m \sum_n a_j^A(x+m,y+n) a_j^B(x+m,y+n) / E^A(x,y) + E^B(x,y) \quad (7)$$

Where $a_j^x(x,y)$; x=A, B is the lowpass contourlet coefficients of the source image A or B and $M_j^{AB}(x,y)$ is the salience factor.

Similarity of the lowpass subbands of the two source images are reflects by the salience factor. After that, this value is compared to a predefined threshold T_L ($T_L = 1$).The averaging mode is chosen for the condition $M_j^{AB}(x,y) > T_L$ with the fusion rules denoted in (8).

$$a_j^F(x,y) = \alpha_A . a_j^A(x,y) + \alpha_B . a_j^B \quad (8)$$

Where $a_j^F(x,y)$ denotes the fused result at position (x,y). The α_A and α_B are selected based on the specified condition, as in (9).

$$\begin{aligned} \alpha_A &= \alpha_{\min} \text{ for } E^A(x,y) < E^B(x,y) \\ &= \alpha_{\max} \text{ for } E^A(x,y) \geq E^B(x,y) \end{aligned} \quad (9)$$

Where $\alpha_B = 1 - \alpha_{A,} \alpha_{\min} \in (0,1) \alpha_{\min} + \alpha_{\max} = 1$

If $M_j^{AB}(x,y) \leq T_L$, selection mode is selected for fusion in (10).

$$\begin{aligned} a_j^F(x,y) &= a_j^A(x,y) \text{ for } E^A(x,y) \geq E^B(x,y) \\ &= a_j^B(x,y) \text{ for } E^A(x,y) < E^B(x,y) \end{aligned} \quad (10)$$

b) Highpass subband fusion:

The coefficients which have larger absolute values in the high frequency subbands $d_{j,k}$ are fused using the average method is defined as follows

$$E_{j,k}^F(x,y) = d_{j,k}^A(x,y) + d_{j,k}^B(x,y) \quad (11)$$

Where $E_{j,k}^F(x,y)$ denotes the local energy, $d_{j,k}^X(x,y)$ is the high frequency coefficient.

c) Reconstruction of fusion image:

The fused image is obtained from $a_j^F(x,y)$ and $E_{j,k}^F(x,y)$ using inverse contourlet decomposition method.

D. Dual-Tree Complex Wavelet Transform

To overcome the drawbacks of DWT, in 1998 kingsbury proposed the dual-tree complex wavelet transform (DTCWT), which provides both good shift variance and directional selectivity. The DTCWT design is based on the use of two parallel trees, first one for odd samples and the second one for the even samples generated at the first stage. These trees provide the signal delays necessary for every level and consequence eliminate aliasing effect. The subbands are divided into six distinct

subbands which are ±15, ±45,±75.

A. Perceptual Fusion Framework for DTCWT

The calculation of JND model proposed by Liu et al. [8] where the local JND associated with a coefficient is defined as:

$$t_{JND}(\lambda,\theta,r,s) = JND_{\lambda,\theta} a_l(\lambda,\theta,r,s) a_c(\lambda,\theta,r,s) \quad (12)$$

Where $JND_{\lambda,\theta}$ the base detection threshold is for DTCWT subband at level λ and orientation θ ; $a_c(\lambda,\theta,r,s)$ is the luminance masking and $a_c(\lambda,\theta,r,s)$ is the contrast masking.

a) Luminance Masking:

The contrast of human visual system is dependent on the luminance context. These model proposed by Chou and Li [9].

$$a_l = \begin{cases} 17\left(1-\sqrt{\dfrac{\bar{I}}{127}}\right)+3, & \text{if } \overline{I \le 127} \\ \dfrac{3}{128}\left(\bar{I}-127\right)+3, & otherwise \end{cases} \quad (13)$$

$\bar{I}$ is calculated from the magnitude of the co-located lowpass coefficient at the highest decomposition level.

a) Contrast Masking:

The contrast masking modelled a_c can be written as:

$$a_c(\lambda,\theta,r,s) = a_{c\_intra}(\lambda,\theta,r,s) a_{c\_inter}(\lambda,\theta,r,s) \quad (14)$$

The intra band contrast masking can be modelled by the nonlinear transducer model introduced by Teo [10], which is written as:

$$a_{c\_intra}(\lambda,\theta,r,s) = \max\left\{1, W_{intra} \sum_{v \in C_{i,j}(h)} \left|\frac{v}{JND_{\lambda,\theta}}\right|^{\varsigma} \Big/ N_{r,s}\right\}, \quad (15)$$

Where (λ,θ,r,s) is the DTCWT coefficient (λ,θ,r,s). The weighting factor W_{intra} and exponent factor ς is set to 12 and 0.6.

The inter band contrast masking written as:

$$a_{c\_inter}(\lambda,\theta,r,s) = \max\left\{1, W_{inter} \sum_{v \in C\lambda,\theta} \omega_\lambda \omega_\theta \left|\frac{v}{JND_{\lambda,\theta}}\right|^{\beta} \Big/ N_{r,s}\right\} \quad (16)$$

Where (λ,θ,r,s) is the DTCWT coefficient (λ,θ,r,s). The weighting factor W_{inter} and exponent factor ς is set to 12 and 0.6.

E. Quantitative Analysis

For the resultant data we are checking the quality of fusion technique by comparing the with various methods by the quantitative measurement. The few metrics explains the quantitative metrics in the following section.

1) Peak Signal to Noise Ratio (PSNR): To measure the quality of the image with respect to the original input image PSNR is used.

$$MSE = \frac{1}{pq}\sum_{i=0}^{p-1}\sum_{j=0}^{q-1}\left[M(i,j)-N(i,j)\right]^2 \quad (17)$$

$$PSNR = 10\log_{10}(MAX^2 / MSE) \quad (18)$$

Where MAX denote the maximum value in an image. p, q are the height and weight of an image. M(i,j) is the input image and N(i,j) is the fused image.

2) Structural Similarity Index (SSIM): SSIM is used to measure the similarity between two images. It is defined as

$$SSIM(A,B) = (2(\mu_A\mu_B + E_1) * 2(\sigma_{AB} + E_2)) / ((\mu_M^2 + \mu_N^2 + E_1) * (\sigma_M^2 + \sigma_N^2 + E_2)) \quad (19)$$

Where μ_M and μ_N is the mean intensities, σ_M and σ_N is the standard deviation, σ_{MN} gives the covariance of M and N, E_1 and E_2 are constants.

3) Mutual Information (MI): let A and B be the two registered images, then the mutual information is given as

$$MI(M,N) = E(M) + E(N) - E(M,N) \quad (20)$$

Where E(M) is the entropy of image M, E(N) is entropy of image N and E(M,N) gives the joint entropy.

III. EXPERIMENTAL ANALYSIS

The following section describes the sequence of the fusion of input images with their performance analysis.

STEP 1: Two given images are taken as test images. Each image have the same pixel size such as 256*256, with 256-level gray scale. Totally, five set of images are used in analyze the performance. One set of input images shown in fig. 6

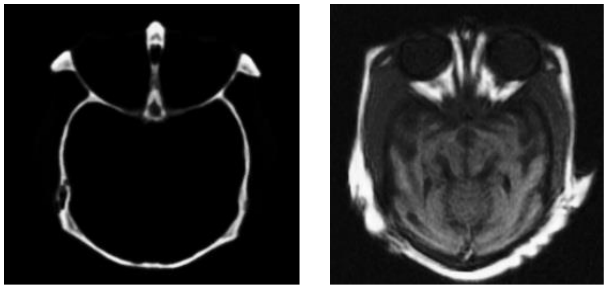


Figure 6. Dataset 1

STEP 2: The input images of each dataset are fused using RTDWT, CT and DTCWT and some of the samples are shown in Fig.7

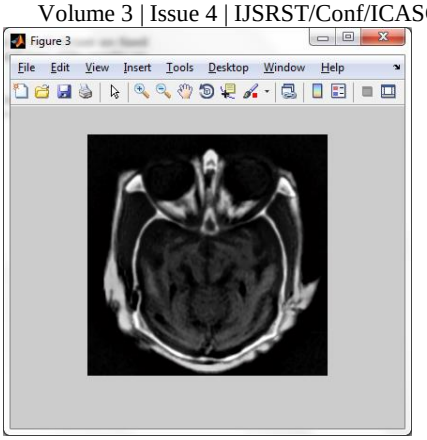


Figure 7. Fused Image of Dataset 1 using DTCWT

STEP 3: Finally, the performance of the fused image is obtained from each technique are analysed with subjective metrics discussed in section II. The evaluation performance of all the techniques, analyzed for the output image with the quantitative metrics are shown in the Table 1

METRICS	ALGORITHM	INPUT SET 1	INPUT SET 2	INPUT SET 3
EN	RDWT 1	6.6693	6.6571	6.5767
	RDWT 2	6.7634	6.7412	6.6638
	RDWT 3	6.8405	6.7866	6.7193
	CT	6.3936	6.1643	6.3261
	DTCWT			
SSIM	RDWT 1	54.6854	53.0163	54.8240
	RDWT 2	59.4352	58.6768	60.1467
	RDWT 3	60.7343	59.0563	60.0563
	CT	56.6345	55.6450	56.6890
	DTCWT	64.7320	66.7854	70.6995
MI	RDWT 1	60.7703	60.0502	60.9051
	RDWT 2	60.4853	59.8634	60.5467
	RDWT 3	60.1785	59.2891	60.2885
	CT	59.1723	59.3468	59.4563
	DTCWT	62.9325	63.0753	62.2335

IV. CONCLUSION

In this paper, by using petrovic dataset we have analyzed the three different fusion techniques using quantitative metrics.

The proposed work has addressed by producing a principle model of the perceptual importance of coefficients within image fusion and then evaluating the performance quantitatively and qualitatively across a representative dataset.

V. REFERENCES

[1]. Firooz Sadjadi, "Comparative Image Fusion Analysis", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 3, June 2005

[2]. Huaixin Chen, "A Multiresolution Image Fusion Based on Principle Component Analysis", Fourth International Conference on Image and Grapics, pp. 737-741, August 2007.

[3]. Oliver Rockinger, "Image Sequence Fusion Using a Shift-Invariant Wavelet Transform", International Conference on Image Processing, Vol. 3, pp.288, October 1997.

[4]. Peter J.Burt, Edward H.Adelson, "The Laplacian Pyramid ass a Compacct Image Code", IEEE Transcations on Communications, Vol.31, pp. 532-540, April 1983.

[5]. James E.Fowler, "The Redundant Discrete Wavelet Transform and Additive Noise", IEEE Signal Processing Letters, Vol.12, No.9, September 2005

[6]. Souparnika Jadhav, "Image Fusion Based on Wavelet Transform", International Journal of Engineering Reasearch, Vol. 3, pp. 442-445, July 2014.

[7]. L.yang, B.L.Guo, W.Ni, "Multimodality Medical Image Fusion Based on Multiscale Geometric Analysis of Contourlet Transform", Elsevier Science Publishers, Vol. 72, pp. 203-211, December 2008.

[8]. Z.liu, L.J.Karam, and A.B.Watson, "JPEG2000 encodind with perceptual distortion control," IEEE Transactions on Image Processing, Vol. 15, no.7, pp. 1763-1778, July 2006.

[9]. C.-H. Chou and Y.-C. Li, "A Perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 5, n0. 6, pp.467-476, 1995.

[10]. P.C.Teo and D.J.Heeger, "Perceptual image distortion", in IEEE International conference on image processing, 1994, pp. 982-986.

[11]. V.Petrovic, "Subjective tests for image fusion evaluation and objective metric validation", Information Fusion, Vol. 8, no. 2, pp. 208-216, 2007