

MULTIFOCUS IMAGE FUSION USING MULTIRESOLUTION APPROACH WITH BILATERAL GRADIENT BASED SHARPNESS CRITERION

G Geetha¹, S.Raja Mohammad², Dr. Y.S.S.R. Murthy³

¹ II year M.Tech, Computer Science & Engineering Department,
Sri Vishnu Engineering College for Women, Bhimavaram, A.P, India.
geetaguttikonda@gmail.com

² Associate Professor, Computer Science & Engineering Department,
Sri Vishnu Engineering College for Women, Bhimavaram, A.P, India.
raja_delip@gmail.com

³ Professor, Computer Science & Engineering Department,
Sri Vishnu Engineering College for Women, Bhimavaram, A.P, India.
yssrmoorthy@gmail.com

ABSTRACT

The fusion of two or more images is required for images captured using different sensors, different modalities or different camera settings to produce the image which is more suitable for computer processing and human visual perception. The optical lenses in the cameras are having limited depth of focus so it is not possible to acquire an image that contains all the objects in-focus. In this case we need a Multifocus image fusion technique to create a single image where all objects are in-focus by combining relevant information in the two or more images. As the sharp images contain more information than blurred images image sharpness will be taken as one of the relevant information in framing the fusion rule. Many existing algorithms use contrast or high local energy as a measure of local sharpness (relevant information). In practice particularly in multimodal image fusion this assumption is not true. Here in this paper we are proposing the method which combines the multiresolution transform and local phase coherence measure to measure the sharpness in the images. The performance of the fusion process was evaluated with mutual information, edge-association and spatial frequency as quality metrics and compared with Laplacian pyramid, DWT (Discrete Wavelet Transform) and bilateral gradient based sharpness criterion methods etc. The results showed that the proposed algorithm is performing better than the existing ones.

KEYWORDS

Multiresolution, Laplacian pyramid, DWT (Discrete Wavelet Transform), Bilateral gradient sharpness criterion.

1. INTRODUCTION

Image fusion is the process of integrating two or more images to form one image which is highly informative. The main aim of this process is to extract the features which are perceptually important from all the images and combine them to form a fused image which is more suitable for computer processing and human visual perception. Image fusion is playing a major role in the

research areas such as Medical Imaging, Remote sensing, computer vision, Robotics and Microscopic imaging etc[1][2].

In the process of image acquisition the image quality depends on the focal length or focus of the optical system. If the lens focusing is poor the resultant image will suffer from blurring. Also it is not possible to focus all the objects in a scene equally. Therefore Multifocus fusion techniques are essential to create an image which contains all the objects are in-focus from two or more images.

Sharp images contain more details than blurred images. In many situations great variations are present in scene's depth; it is difficult to acquire an image in which all the areas of scene appear sharp. The scene areas which are in-focus will appear sharp and has higher contrast. Areas which are out of focus i.e. in front of or behind the focus plane will be blurred. So the key challenge of image fusion process is to identify the relevant information such as contrast/ high local energy and areas that are having high sharpness from the best-focused images and combine this useful information to create a highly informative image [1][2].

In this paper we are using multiresolution transform along with the gradient information of the images to measure the strength and phase coherence which are the deciding factors in measuring sharpness of the image. An image with higher phase coherence and higher strength is considered as sharper and more informative.

The paper is organized as follows section 2 presents the formulation of Multifocus image fusion problem and reviews various sharpness measures existed in the literature. Section 3 introduces the multiresolution transforms used in this paper and section 4&5 explains the proposed method for image fusion. Section 6 will present the results and compare them with the existing fusion techniques finally section 7 concludes the paper.

2. PROBLEM FORMULATION

Given a set of 2-D images $I_1(x, y), I_2(x, y), \dots, I_n(x, y)$ which are captured with different imaging settings in the camera and aligned well, the aim of multi-focus image fusion is to combine the information content of the individual images into a single fused image $f(x, y)$. Normalized aggregation of the images is a simple method for image fusion This can be expressed as

$$f(x, y) = \frac{1}{N} \sum_{n=1}^N I_n(x, y) \quad (1)$$

The main drawback with this method is that all information content within the images is given equal importance. Therefore, image regions such as edges and boundaries which are more informative, are treated no differently than unimportant regions. A normalized weighted aggregation approach was introduced to overcome this situation. It can be mathematically expressed as

$$f(x, y) = \frac{\sum_{n=1}^N w_n(x, y) I_n(x, y)}{\sum_{n=1}^N w_n(x, y)} \quad (2)$$

where $w_n(x, y)$ is the weight assigned to information content at (x, y) in the n-th image. The choice of the weighting scheme (i.e., the determination of $w_n(x, y)$ in Eq. (2) is important in the performance of the image fusion algorithm which is required to be locally adaptive to image's content and be sensitive to the blur of each image. In view of this, a brief survey on existing sharpness measures will be discussed here and bilateral sharpness criterion using two kinds of statistics of image's gradient information in Section 4.

2.1. Existing Sharpness Criteria

To measure the degree of images blur the high-frequency information such as edges and boundaries are usually used as the basis. In general sharper edges are present in the well-focused images and is expected to have higher frequency content than those that are blurred. In the following analysis, denote $I(x, y)$ be the intensity value at the position (x, y) of the image I [3][4].

Variance [3]. For an $M \times N$ block of the image, its variance is defined as

$$S_{VAR} = \frac{1}{M \times N} \sum_x \sum_y (I(x, y) - \mu)^2 \quad (3)$$

Where μ is the mean intensity value of the image block and it is defined as

$$\mu = \frac{1}{M \times N} \sum_x \sum_y (I(x, y)) \quad (4)$$

Energy of image gradient [3]. For a $M \times N$ block of the image, it is measured as [12]

$$S_{EG} = \sum_x \sum_y (I_x^2 + I_y^2) \quad (5)$$

Where I_x and I_y represent image gradients at the row and column directions, respectively. They are usually defined as $I_x = I(x+1, y) - I(x, y)$ and $I_y = I(x, y+1) - I(x, y)$.

Tenenbaum. For an block of the image, it is measured as

$$S_{TNG} = \sum_x \sum_y (\nabla I(x, y))^2 \quad (6)$$

Where $\nabla I(x, y) = \sqrt{I_x^2 + I_y^2}$ in which I_x and I_y are gradients

(obtained using Sobel operators) along the row and column directions, respectively.

Energy of Laplacian [3]. For an $M \times N$ block of the image, it is measured as

$$S_{EL} = \sum_x \sum_y (\nabla^2 I(x, y))^2 \quad (7)$$

where $\nabla^2 I(r, c)$ represents image gradient obtained by Laplacian operator $[-1, -4, -1; -4, 20, -4; -1, -4, -1]$.

Sum-modified-Laplacian [3]. It differs from the usual Laplacian operator in that the absolute values of the partial second derivatives are summed instead of their actual values. That is, it is mathematically expressed as

$$S_{SML} = \sum_x \sum_y (\nabla^2 I(x, y))^2 \quad (8)$$

Where $\nabla^2 I(x, y) = |2I(x, y) - I(x+1, y) - I(x-1, y)| + |2I(x, y) - I(x, y+1) - I(x, y-1)|$.

Frequency selective weighted median filter [14]. It measures the sharpness of the image as

$$S_{FSWM} = \sum_x \sum_y (I_x^2 + I_y^2) \quad (9)$$

Where

$$\begin{aligned}
I_y &= \text{med}\{I(x, y-1), I(x, y), I(x, y+1)\} - \frac{1}{2} \text{med}\{I(x, y-3), I(x, y-2), I(x, y-1)\} \\
&\quad - \frac{1}{2} \text{med}\{I(x, y+1), I(x, y+2), I(x, y+3)\}. \\
I_x &= \text{med}\{I(x-1, y), I(x, y), I(x+1, y)\} - \frac{1}{2} \text{med}\{I(x-3, y), I(x-2, y), I(x-1, y)\} \\
&\quad - \frac{1}{2} \text{med}\{I(x+1, y), I(x+2, y), I(x+3, y)\},
\end{aligned}$$

Phase coherence model [9][11][13]: It is consistent to the perceptual significance of the image and it can be determined at a particular position (x,y) as

$$\begin{aligned}
S_{PCM}(x, y) &= \frac{1}{2} \sum_{\theta} \left[(h(x, y, \theta) \sin(\theta))^2 + (h(x, y, \theta) \cos(\theta))^2 \right] \\
&\quad + \frac{1}{2} \sqrt{4 \left(\sum_{\theta} (h(x, y, \theta) \sin(\theta))^2 + (h(x, y, \theta) \cos(\theta))^2 \right) + \left(\sum_{\theta} [(h(x, y, \theta) \cos(\theta))^2 - (h(x, y, \theta) \sin(\theta))^2] \right)^2},
\end{aligned} \tag{10}$$

$$h(x, y, \theta) = \frac{\sum_n w(x, y, \theta) |A_n(x, y, \theta) \nabla \varphi_n(x, y, \theta)|}{\sum_n A_n(x, y, \theta) + \xi},$$

Where

$$\Delta \varphi_n(x, y, \theta) = \cos(\varphi_n(x, y, \theta) - \bar{\varphi}_n(x, y, \theta)) - \left| \sin(\varphi_n(x, y, \theta) - \bar{\varphi}_n(x, y, \theta)) \right|$$

in which W represents the frequency spread weighting factor, A_n and φ_n represent the amplitude and phase at the wavelet scale n, respectively, $\bar{\varphi}_n$ represents the weighted mean phase, ξ is a small constant used to avoid the division by zero. All of these parameters are as same as that used in.

3. MULTIREOLUTION APPROACH

3.1. Laplacian Pyramid

Several signal processing methods for signal enhancement, compression, denoising, and fusion etc. were developed based on Laplacian pyramid since Burt and Adelson introduced this transform back in 1983[5]. The Laplacian Pyramid implements a “feature (pattern) selective” approach to image fusion, in this method the fused image is constructed not a pixel at a time, but a feature at a time.

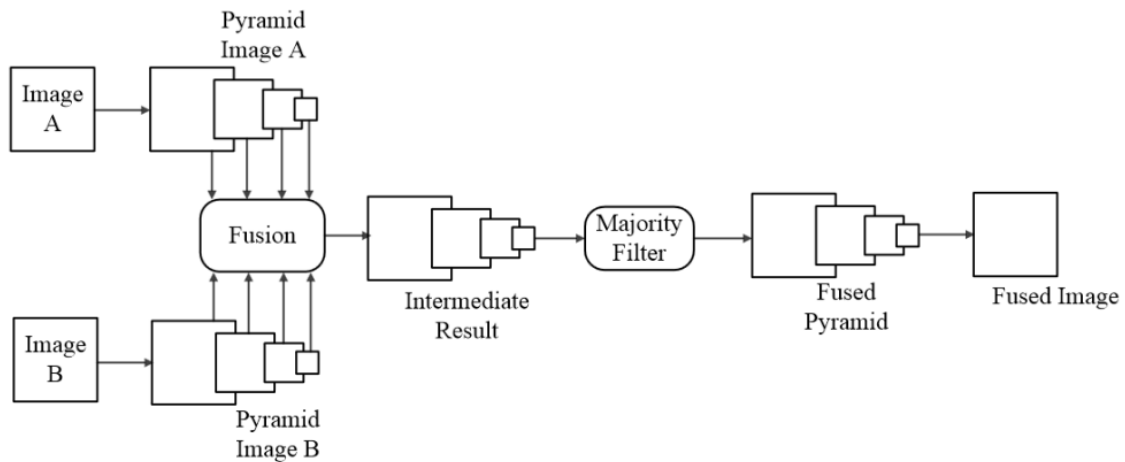


Figure 1: Image fusion system using Laplacian Pyramid

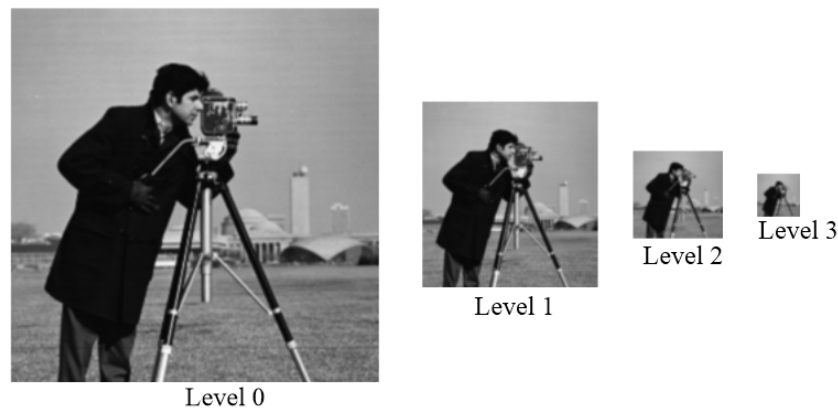


Figure 2: Images from Laplacian Pyramid

In multi resolution pyramid transformation an image is decomposed into multiple resolutions at different scales. A pyramid is a sequence of images in which each level is a filtered and subsampled copy of the predecessor. Higher levels of the pyramid are increased scale versions and reduced resolution of the original image. The lowest level of the pyramid has the same scale as the original image and contains the highest resolution information. Figure 1 shows a schematic of pyramid multi-resolution representation [5][6][7].

Step 1: Verify all the Images are of same size or not if not reject the fusion process.

Step2: Decompose the Images into multiple resolutions at different scales.

Step3: Do the fusion process scale by scale by applying the appropriate fusion rule (Ex: selecting the maximum coefficients.)

Step4: Reconstruct the image from the resultant scales to get the fused image.

3.2. DWT (Discrete Wavelet Transform)

The DWT of a signal $x(n)$ is calculated by passing the signal through a series of filters. The samples are passed through a low pass filter first with impulse response $g(n)$ resulting in a convolution of the two [5][7]:

$$y[n] = (x * g)[n] = \sum_{k=-\infty}^{\infty} x(k)g(n-k) \quad (11)$$

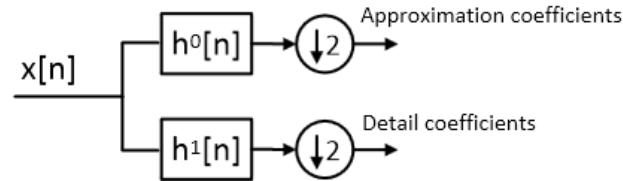


Figure 3: DWT Single Stage Decomposition

Simultaneously the signal is also decomposed using a high-pass filter h . The outputs of the highpass filter are called detail coefficients and the outputs of the lowpass filter are called approximation coefficients. The two filters i.e lowpass and highpass are related to each other and are known as QMF (quadrature mirror filter) pair. In this process half the frequencies of the signal have been removed, so half the samples can be discarded according to Nyquist's rule. The filter outputs are then subsampled by 2.

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x(k)g(2n-k) \quad (12)$$

$$y_{high}[n] = \sum_{k=-\infty}^{\infty} x(k)h(2n+1-k)$$

This decomposition has halved the time resolution since only half of each filter output characterizes the signal. So the frequency resolution has been doubled since each output has half the frequency band of the input

2D DWT can be implemented by applying the 1D DWT along the rows of an image first and applying then on the columns of an image. When a wavelet transform is applied to an image which is a 2D signal it decomposes the image into four subbands as shown in the "fig.4". The LL band contains the approximation coefficients, LH band contains horizontal details, HL band contains vertical details and HH band will contain the diagonal details.

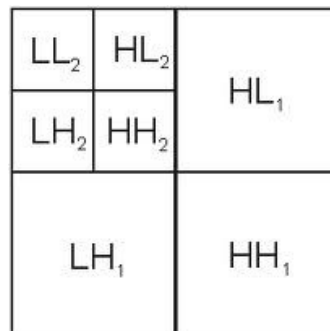


Figure 4: 2D DWT Decomposition of Image into subbands using Discrete Wavelet Transform

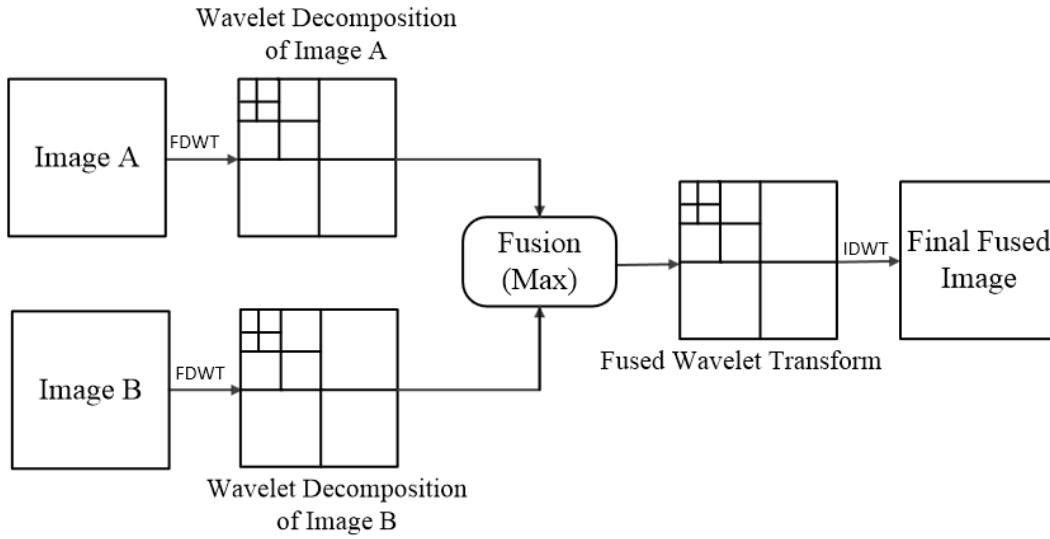


Figure 5: Image Fusion system using Discrete Wavelet Transform

Step 1: Verify all the Images are of same size or not if not reject the fusion process.

Step2: Decompose the Images into multiple resolutions at different scales.

Step3: Do the fusion process subband by subband by applying the appropriate fusion rule (Ex: selecting the maximum coefficients.)

Step4: Reconstruct the image from the resultant subbands to get the fused image.

4. BILATERAL GRADIENT BASED SHARPNESS CRITERION

The statistics of image's gradient is examined in this section to propose a new sharpness measurement criterion, which exploits bilateral statistics of image's gradient information [9][11][13]. Using the image gradients Image structure can be measured effectively. Consider an image of interest $I(x, y)$. The gradient covariance matrix of a region within an $M \times N$ local window is defined as

$$C = \begin{pmatrix} \sum_w I_x^2(x, y) & \sum_w I_x(x, y)I_y(x, y) \\ \sum_w I_x(x, y)I_y(x, y) & \sum_w I_y^2(x, y) \end{pmatrix} \quad (13)$$

Where $I_x(x, y)$ and $I_y(x, y)$ represent image's gradient at the row and column directions, respectively. The above gradient covariance matrix can be decomposed as

$$C = VDV^T = (v_1 \ v_2) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \end{pmatrix} \quad (14)$$

where V represents a 2×2 matrix whose column vectors are eigenvectors v_1 and v_2 , D denotes the 2×2 diagonal matrix whose diagonal elements are eigenvalues λ_1 and λ_2 ($\lambda_1 > \lambda_2$) that correspond to eigenvectors v_1 and v_2 , respectively, and the superscript T denotes the transpose.

The geometrical structure at a pixel in an image can be described by the eigenvalues λ_1 and λ_2 of the above gradient covariance matrix. Motivated by this, the first criterion is proposed to measure the strength of the image's gradient, which is defined as

$$A(x, y) = \lambda_1 - \lambda_2 \quad (15)$$

On the other hand, local phase coherence is consistent with the perceptual significance of image's characteristics. This has been supported by the physiological evidence that showed high human perception response to signal characteristics with high local phase coherence. Another advantage is the fact that it is insensitive magnitude variations caused by illumination conditions or noise present in image signals. In view of this, the second criterion is to consider the phase coherence of the image's gradient, that is

$$P(x, y) = -\cos(\theta(x, y) - \overline{\theta}(x, y)) \quad (16)$$

where $\theta(r, c)$ is the phase information at the position (r, c) determined by the principle Eigen vector v_1 associated with the largest eigenvalue λ_1 defined as (13), $\overline{\theta}_{r,c}$ is the average of phases of the neighbouring positions. This measure achieves the maximal value when the local phase coherence is worst, which is usually caused by an edge. Finally, the above two criteria (15) and (16) are jointly considered to develop a bilateral sharpness criterion as

$$S_{BSC} = A^\alpha(x, y)P^\beta(x, y) \quad (17)$$

Where α and β are two factors to adjust contributions of two criteria.

5. FUSION PROCESS

Step 1: Verify all the Images are of same size or not if not reject the fusion process.

Step2: Decompose the Images into multiple resolutions at different scales by using multiresolution transform

Step3: Do the fusion process subband by subband by applying the bilateral gradient based sharpness criterion.

Step4: Reconstruct the image from the resultant subbands to get the fused image.

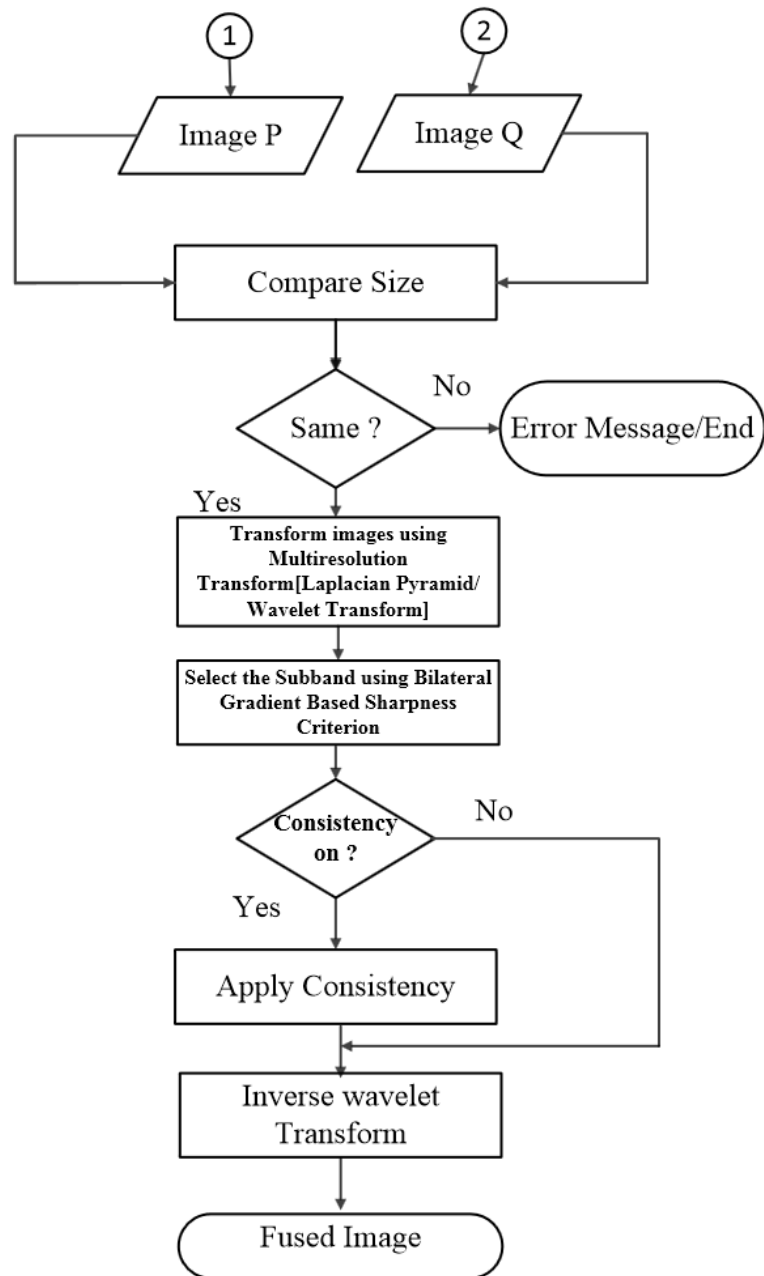


Figure 6: Flowchart of the proposed fusion process

6. RESULTS

To evaluate the performance of the fusion process in this paper we used the Mutual Information (MI), edge-association (QAB/F), spatial frequency (SF) along with the RMSE (Root Mean Square error) and STD (Standard Deviation). The comparison of the proposed algorithm was carried out with the existing Bilateral Gradient based Sharpness criterion, Laplacian Pyramid method and Image fusion using DWT.

RMSE:

RMSE will measure the difference between the ideal image and the fused image. It is defined as

$$RMSE = \sqrt{\frac{1}{MN} \sum_{x=0}^M \sum_{y=0}^N [g(x, y) - f(x, y)]^2} \quad (18)$$

Spatial Frequency (SF):

Spatial frequency is defined as

$$SF = \sqrt{RF^2 + CF^2} \quad (19)$$

Where RF and CF are the row frequency

$$RF = \sqrt{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - f(x, y-1)]^2}$$

and column frequency

$$CF = \sqrt{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - f(x-1, y)]^2}$$

respectively and $f(x, y)$ is the fused image.

Mutual Information (MI)

We know that in information theory Mutual Information is a basic concept for measuring the statistical dependence between two random variables and the amount of information that one variable contains about the others. Mutual Information measure here describes the similarity of the image intensity distributions of the corresponding image pair [14][16].

Let A and B be two random variables with marginal probability distributions $p_A(a)$ and $p_B(b)$ and joint probability distribution $p_{AB}(a, b)$. MI measures the degree of dependence of the two random variables A and B . It is defined by Kullback-Leibler measure

$$I_{AB}(a, b) = \sum_{x, y} p_{AB}(a, b) \log \frac{p_{AB}(a, b)}{p_A(a)p_B(b)} \quad (20)$$

The fused image will contain the important information from the original input image sets. Here 'important information' depends on the application and is difficult to define. Mutual information is the amount of information that one image contains about another.

Considering two input images A, B and a fused image F we can calculate the amount of information that F contains about A and B according to above equation

$$I_{FA}(f, a) = \sum_{x, y} p_{FA}(f, a) \log \frac{p_{FA}(f, a)}{p_F(f)p_A(a)}$$

$$I_{FB}(f, b) = \sum_{x, y} p_{FB}(f, b) \log \frac{p_{FB}(f, b)}{p_F(f)p_B(b)}$$

Thus the mutual information is given as

$$M_F^{AB} = I_{FA}(f, a) + I_{FB}(f, b) \quad (21)$$

The above equation indicates that the total information that a fused image F contain about A and B.

Edge Association

In this measure an important visual information with the “edge” information that is present in each pixel of an image. The visual to edge information association is well supported by HVS (Human Visual System) studies and is frequently used in compression systems and image analysis [14][15][16].

The amount of edge information that is transferred from input images to the fused image can be obtained as

$$Q_p^{AB/F} = \frac{\sum_{x=1}^M \sum_{y=1}^N Q^{AF}(x, y)w^A(x, y) + Q^{BF}(x, y)w^B(x, y)}{\sum_{x=1}^M \sum_{y=1}^N (w^A(x, y) + w^B(x, y))} \quad (22)$$

Where Q^{AF}, Q^{BF} are edge preservation values and w^A, w^B are the corresponding weights

Table 1: Performance evaluation of various Fusion Algorithms

Fusion Method	MI	QAB/F	Spatial Frequency
Average	4.9722	0.6629	09.1078
Select Maximum	5.6043	0.5793	09.2172
Select Minimum	5.9159	0.7340	12.4169
Laplacian Pyramid	4.5623	0.8041	16.9250
Gradient Pyramid	3.9768	0.7570	13.7894
DWT	4.1178	0.7930	16.7954
Bilateral Gradient	5.3671	0.8020	16.9230
Laplacian Pyramid +Bilateral Gradient	6.4671	0.8302	17.6230
DWT + Bilateral Gradient	6.9671	0.8920	18.4230



Figure 7: Input images for the fusion process (a) Bottom object is in-focus (b) TOP object is in-focus



Figure 8: (a) Fusion based on bilateral gradient based sharpness criterion (b) Fusion using DWT



Figure 9: (a) Fusion using Laplacian Pyramid + bilateral gradient based sharpness criterion (b) Fusion using DWT + bilateral gradient based sharpness criterion

7. CONCLUSIONS AND FUTURE WORK

The above results shows that the DWT based image fusion is introducing little amount of blur in the fused image while the combination of bilateral gradient with Laplacian pyramid and DWT are giving better fused images than the existing ones. The performance of the algorithms can be improved by introducing the directional oriented multiresolution transforms such as steerable pyramids, contourlets etc and shift invariant transforms such as Undecimated wavelet transforms and complex wavelet transforms in the multiresolution decomposition stage.

REFERENCES

- [1] A.A. Goshtasby, S. Nikolov, *Information Fusion* 8 (Apr. 2007) 114.
- [2] H.B. Mitchell, *Image fusion: theories, techniques and applications*, Springer, 2010.
- [3] W. Huang, Z. Jing, "Evaluation of focus measures in multi-focus image fusion" *Pattern Recognition Letters* 28 (Apr.2007) 493.
- [4] Chong-Yaw Wee, Measure of image sharpness using Eigen values, *Information Sciences an International Journal*,177(2007), 2533-2552.
- [5] G. Pajares, J.M. Cruz, A wavelet-based image fusion tutorial, *Pattern Recognition* vol. 37 (Sept.2004) 1855.
- [6] Shutao Li, Bin Yang, Multifocus image fusion by combining curvelet and wavelet transform, *Pattern Recognition Letters*29(2008), 1295-1301.
- [7] Jing Tian, Li Chen, Multi-Focus image fusion using wavelet statistics, *Proceedings of IEEE 17th conference on Image Processing*, September 26-29, Hongkong, 1205-1208.
- [8] S. Arivazhagan, L. Ganesan, T.G. Subash Kumar, A modified statistical approach for image fusion using wavelet transform *Signal, Image and Video Processing* 3 (Jun. 2009) 137.
- [9] R. Hassen, Z. Wang, M. Salama, Multifocus image fusion using local phase coherence measurement, *Proc. Int. Conf. on Image Analysis and Recognition*, Jul.2009, p. 54.
- [10] Jing Tian, Li chen, Multifocus image fusion using a bilateral gradient based sharpness criterion, *Optics communications* 284 (2011), 80-87
- [11] Alexander Wong, William bishop, Efficient least squares fusion of MRI and CT images using a phase congruency model, *Pattern Recognition letters* 29 (2008), 173-180.
- [12] Yingjie Zhang, Liling Ge, Efficient fusion scheme for multi- focus images by using blurring measure, *Digial Signal Processing* 19 , (2009), 186-193.
- [13] A. Agrawal, R. Raskar, R. Chellappa, Edge suppression by gradient field transformation using cross-projection tensors, *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 2006, p. 2301.
- [14] V. Aslantas, R. Kurban, Evaluation of criterion functions for the fusion of Multifocus noisy images, *Proc. Int. Conf. on Signal Processing and Communications Applications*, Apr.2009, p. 492.
- [15] Shutao Li, James T kwok, Combination of images with diverse focuses using the spatial frequency, *Information Fusion* 2 (2001), 169-176.
- [16] P.W. Huang, C.-I. Chen, P.-L. Lin, Multi-focus image fusion based on salient edge information within adaptive focus- measuring windows, *Proc. Int. Conf. on Systems, Man and Cybernetics*, Oct. 2009, p. 2589.