

# Multi-Focus Image Fusion in Transform Domain using Steerable Pyramids

Kiranjeet Kaur

Student

Ludhiana College of Engineering &  
Technology Katani Kalan

Gurinder Singh

Asst. Professor

Department of Electronics &  
Communication Engineering  
Ludhiana College of Engineering &  
Technology Katani Kalan

## ABSTRACT

In optical lenses of conventional cameras, depth of field is restricted to a particular range. Therefore, only those objects that are at a particular distance from the camera captured clearly with proper focus whereas objects at other distances in front of or behind the focus plane remain defocused and blurred. However, for accurately interpreting and analyzing images, it is desired to obtain images with every object in focus. Multi-focus image fusion is an effective technique to solve this problem by combining two or more images of the same scene taken with different focus settings into a single all-in-focus image with extended depth of field, which is very useful for human or machine perception. The main drawbacks of pixel based fusion methods are misalignment of decision map with boundary of focused objects and wrong decision in sub-regions of the focused or defocused regions which produce undesirable artifacts in the final fused image. Therefore frequency domain methods are more preferred than spatial domain methods. In previous years, many kinds of multi-scale transforms have been proposed and adopted for image fusion such as pyramid decomposition, discrete wavelet transform (DWT), dual-tree complex wavelet transform (DTCWT), and discrete cosine harmonic wavelet transform (DCHWT). These transforms are widely used as wavelets become dominant filters in most of the techniques but they still have drawbacks. Wavelets are the lack of translation invariance, especially in two-dimensional (2D) signals and the poor selectivity in orientation. This can be overcome by steerable pyramid transform as one can choose the orientation before applying the filters. We have applied this method to achieve multi-focus image fusion of images which have different focus areas while capturing them. Experimental results show that the proposed method affectively carried out fusion process as performance of the technique has been evaluated by various parameters namely mutual information, QABF factor used for edge perservance measuring and entropy etc.

## Keywords

Image fusion, gray scale images, multi-focus, pyramid transforms, wavelets

## 1. INTRODUCTION

As depth of field in optical lenses of conventional cameras is limited, the lens of a camera can only focus at one distance or on a plane at a time. Thereby, those objects that fall in the depth of field (DOF) of the lens appear sharp, while the objects out of the DOF are blurred. As a result, the multiple objects of interest may not be all focused in one single image. However, for accurately interpreting and analyzing images, it is desired to obtain images with every object in focus [1]. Multi-focus image fusion is an effective technique to solve this problem by combining two or more images of the same

scene taken with different focus settings into a single all-in-focus image with extended depth of field, which is very useful for human or machine perception. The multi-focus image fusion has been applied in various applications such as microscopic imaging, remote sensing, and computer vision [1].

## 1.1 Spatial And Transform Domain Method Of Image Fusion

In the past decades, different genres of multi-focus image fusion algorithms have been proposed. Most of these algorithms are done either in the spatial domain or transform domain. The transform domain algorithms usually extract coefficients with respect to some localized bases [2]. Then, these coefficients are processed according to some criteria and transformed back to produce the fusion image. The spatial domain algorithms are based on a local operation. They can avoid the global problems and preserve more original information from the source images than the transform domain algorithms but in many other factors, transform domain methods are better i.e. calculation time and appropriate decision making in selection rules etc. In general, spatial domain algorithms can be classified as: pixel-based, block-based, and region-based [3]. The pixel based fusion algorithms just consider the single pixels or use the information in the local neighborhoods, which could lead to some undesirable side effects [2], such as reduced contrast, and blocking artifacts. To resolve this problem, some block-based algorithms have been proposed. In the block-based algorithms, the source images are firstly decomposed into blocks with equal size. Then the focused blocks can be detected by measuring the focus on the corresponding blocks. However, a large block may contain both the focused and defocused regions and the focus of a small block could not be well measured [3]. Therefore, the performance of these algorithms is restricted by the selected block-size.

## 1.2 Limitations or drawbacks in image fusion methods and fusion approaches

The main drawbacks of pixel based fusion methods are misalignment of decision map with boundary of focused objects and wrong decision in sub-regions of the focused or defocused regions which produce undesirable artifacts in the final fused image. Therefore frequency domain methods are more preferred than spatial domain methods. In the first group, transform coefficients are fused and the fused image is reconstructed from these composite coefficients. The transform domain fusion methods that are based on multi-scale transforms are the most commonly used methods. Many kinds of multi-scale transforms have been proposed and adopted for image fusion such as pyramid decomposition, discrete wavelet transform (DWT), dual-tree complex

wavelet transform (DTCWT), and discrete cosine harmonic wavelet transform (DCHWT). The second main factor other than choice of spatial and frequency methods is the selection of fusion rules. The basic assumption of the multi-focus image fusion is that focused objects seem sharper than unfocused ones. So the key issue in the field is how to formulate information measures to detect which parts are in focus, and which parts not. However, the fusion schemes which follow the basic assumption cannot always get satisfactory fusion results, especially in smooth regions. Therefore a suitable method needs to be adopted to take the decision while fusion of the images. In this work we have moved from basic DWT technique which is widely used in literature to a more effective one. We have used steerable pyramid technique (SPT) and various fusion rules are applied to analyse the behavior of this transform. So before defining the proposed method a brief has been given in section below about the previous work done in image fusion

## **2. LITERATURE SURVEY**

In recent years, several more reasonable fusion algorithms have been proposed.

In [4] [5], several region-based algorithms are proposed for multi-focus image fusion. At first, the source images are segmented into regions, and then the focused regions could be found out by comparing the focus-measures of the corresponding regions. However, the region-based algorithms are time-consuming and their performance greatly relies on its previous segmentation results.

In [6], De et al. proposed a quad tree-based algorithm for multi-focus image fusion. In this algorithm, the source images are firstly decomposed into blocks with different sizes in a quad tree structure. In this tree structure, the focused blocks are detected by measuring the focus on the corresponding blocks. But in some cases, their quad tree decomposition method might be not effective, so that the optimal subdivision of the source images could not be achieved.

Prakash et al. [7] propose a pixel-level image fusion scheme using multi resolution Bi-orthogonal wavelet transform (BWT). Two important properties wavelet symmetry and linear phase of BWT have been exploited for image fusion because they are capable to preserve edge information and hence reducing the distortions in the fused image. The performance of the proposed method have been extensively tested on several pairs of multi focus and multimodal images both free from any noise and in presence of additive white Gaussian noise and compared visually and quantitatively against existing spatial domain methods.

Liu et al. [8] proposed an image fusion scheme based on the lifting scheme of wavelets. The pending images are decomposed by using the wavelet lifting scheme into four sub bands firstly: LL, LH, HL, HH, and then sub band LH, HL, HH, are synthesized to obtain three directions of high-

frequency details of the images. The local luminance contrast, which is represented by weighted region energies, is calculated by the Gaussian kernel based on the high-frequency details. Thus the energy-based image fusion rule is applied to get a binary map by choosing the maximum energy between images.

Zhang et al. [9] proposed a medical image fusion based on wavelet theory. In his work, image processing get across multi resolution characteristics of wavelet to de noise, image registration pass the wavelet analysis to gain biggish change point and receive image edge to achieve quick and nice super possible, image fusion use disassemble image to different frequency sub band to save all information to have a perfect fusion

Su et al. [10] proposed a technique in which the IR image and visible image are decomposed by wavelet transform and their multi-scale sub-images. Then, the contrast of IR image is improved by modifying the modulus of the sub-images in scale space and stretching the dynamic scope of smooth sub image at coarser resolution level. Finally, the improved IR images and visible images are fused at different scales and reconstructed to the fused image

## **3. PROPOSED METHOD**

In this work, Steerable pyramid transform or SPT representation has been used as a basic method for decomposition to multi levels. In signal processing, a signal can be decomposed into sub-bands, such as by wavelet transform. The wavelet transform is widely used in many applications since the pyramid structure of wavelets responds well to a human visual system. However, the two major drawbacks of wavelets are the lack of translation invariance, especially in two-dimensional (2D) signals (Cheon and Kim, 2009) and the poor selectivity in orientation. To overcome this problem, the “steerable” pyramid wavelet, a class of arbitrary orientation filters generated by linear combination of a set of basis filters, has been proposed (Simoncelli et al., 1992, 1995). The system diagram of a steerable pyramid for a single stage is shown in Figure 1. The pyramid is divided into two parts: analysis and synthesis. On the analysis part, the image is decomposed into low-pass and high-pass sub-bands, using steerable filters  $L_0$  and  $H_0$ . The low-pass band continues to breakdown into a set of band-pass sub-bands  $B_0$ ;  $B_K$  and lower low-pass sub-band  $L_1$ . The lower low-pass sub-band is sub-sampled by along the two directions  $x$  and  $y$ . Repeating the shaded area provides the recursive structure. Due to its invariant properties, the pyramid structure of the steerable wavelet is used in image fusion. After decomposition, fusion rules has been applied on decomposed coefficients which results in modified coefficients in low frequency as well as band-pass frequency sub-images produced in decomposition process in multi-level pyramid decomposition. The whole process of image fusion has been described in the block diagram below

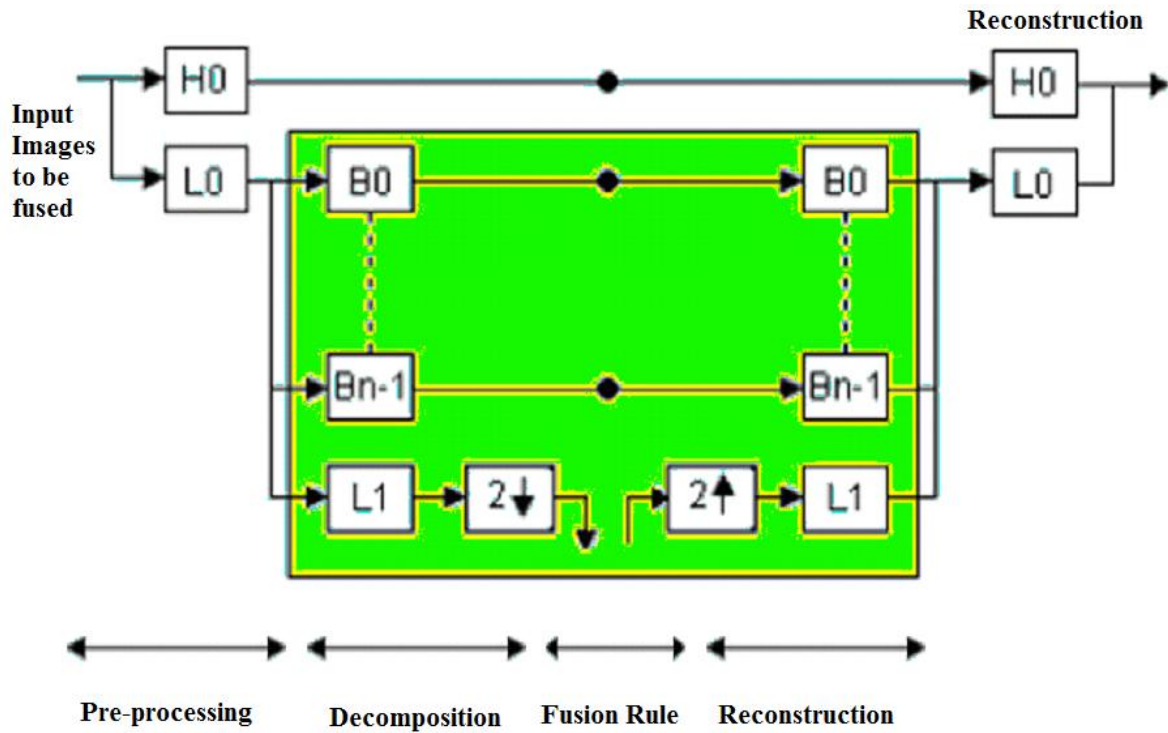


Figure 1: First level of steerable pyramid decomposition system and fusion place of the coefficients of two images.

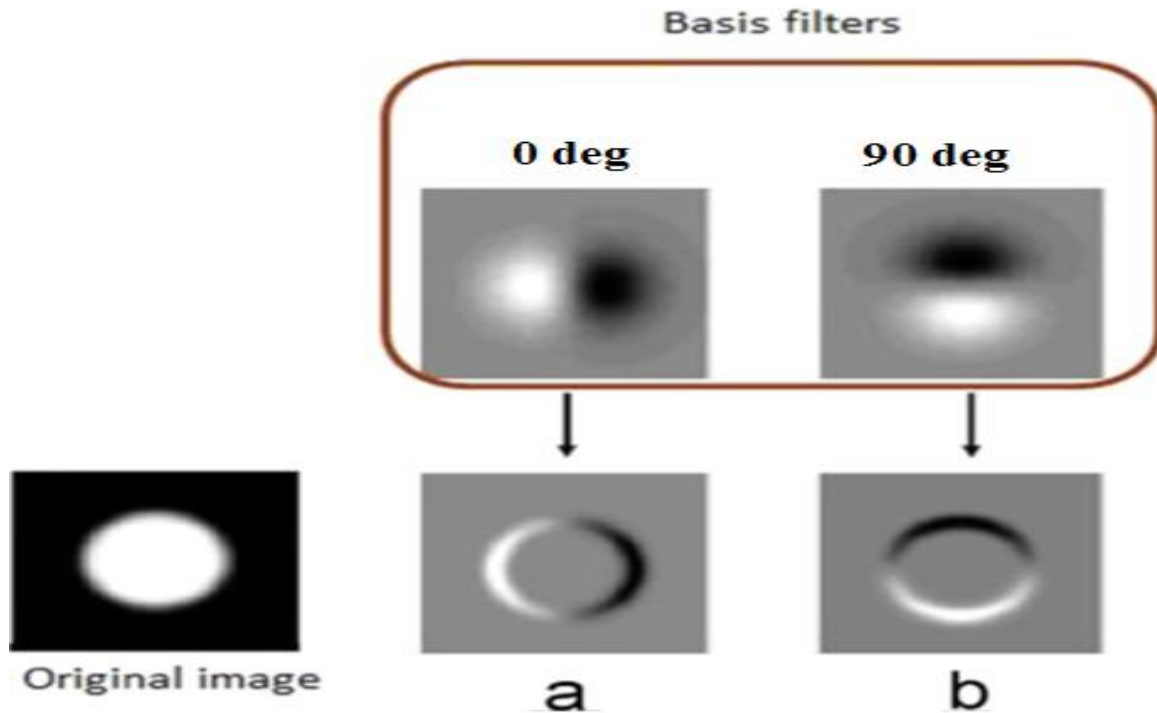


Figure 2: Responses of the circular disk image convolved to the filter :(a) at 0 degree, (b) at 90 degree.

The derivatives of an image in any direction can be interpolated by several basis derivative functions. Any  $K$ th - order directional derivative is a linear combination of  $(K+1)$   $K$ th-order basis derivatives. For example the response of the first-order steered filter in Cartesian co-ordinates  $f_1(x, y)$  to an arbitrary direction  $\theta$  can be easily interpolated and synthesized by taking a linear combination of the basis filters  $f_1^0$  and  $f_1^{90}$  as shown by the following equation:

$$f_1^\theta(x, y) = \cos(\theta)f_1^0(x, y) + \sin(\theta)f_1^{90}(x, y) \quad (1)$$

where  $f_1^0$  and  $f_1^{90}$  are the two first-order basis derivatives given respectively by Eqs. (2) and (3)

$$f_1^0(x, y) = \frac{\partial}{\partial x} e^{-(x^2+y^2)} = -2xe^{-(x^2+y^2)} \quad (2)$$

$$f_1^{90^\circ}(x, y) = \frac{\partial}{\partial x} e^{-(x^2+y^2)} = -2ye^{-(x^2+y^2)} \quad (3)$$

Fig. 2 shows the convolution result of the disk image with the two basis filters  $f_1^0$  and  $f_1^{90}$

Steerable pyramid is best defined in the Fourier domain where it provides a polar-separable decomposition, thus allowing independent representation of scale and orientation. Considering the polar-separability of the filters in the Fourier domain, the first low, and high-pass filters, are defined as

$$L_0(r, \theta) = \begin{cases} 1 & \text{if } r \leq \frac{\pi}{4} \\ \cos\left(\frac{\pi}{2} \log_2\left(\frac{2r}{\pi}\right)\right) & \text{if } \frac{\pi}{4} < r < \frac{\pi}{2} \\ 0 & \text{if } r \geq \frac{\pi}{2} \end{cases} \quad (4)$$

$$H_0(r, \theta) = \begin{cases} 1 & \text{if } r \leq \frac{\pi}{4} \\ \cos\left(\frac{\pi}{2} \log_2\left(\frac{r}{\pi}\right)\right) & \text{if } \frac{\pi}{4} < r < \frac{\pi}{2} \\ 0 & \text{if } r \geq \frac{\pi}{2} \end{cases} \quad (5)$$

where  $r, \theta$  are the polar frequency coordinates. The Fourier magnitude of the  $K$ th oriented band-pass filter is given in polar-separable form by the following equation:

$$B_k(r, \theta) = \text{Ang}_{k(\theta)} \text{Rad}(r)$$

Where  $k = \{0, \dots, K-1\}$ . (6)

$B(r, \theta)$

represents the  $K$  directional band pass filters used in the iterative stages, with radial and angular parts, defined as

$$\text{Rad}(r) = \begin{cases} 1 & \text{if } r \geq \frac{\pi}{4} \\ \cos\left(\frac{\pi}{2} \log_2\left(\frac{2r}{\pi}\right)\right) & \text{if } \frac{\pi}{4} < r < \frac{\pi}{2} \\ 0 & \text{if } r \leq \frac{\pi}{2} \end{cases} \quad (7)$$



Figure 3: A sample of images to be fused

After this fusion rules are applied to the low frequency and high frequency co-efficient using following rules

First Method: Average of low frequency co-efficient Mean of sub-band orientation co-efficient

It is a well-documented fact that regions of images that are in focus tend to be of higher pixel intensity. Thus this algorithm is a simple way of obtaining an output image with all regions in focus. The value of the pixel of each image is taken and added. This sum is then divided by 2 to obtain the average. The average value is assigned to the corresponding pixel of the output image. This is repeated for all pixel values for sub-band and low frequency bands.

Second Method: Minimum of low frequency co-efficient Mean of sub-band orientation co-efficient

The smaller the pixel values the more in focus the image. Thus this rule chooses the in focus regions from each input image by choosing the smallest value for each pixel, resulting in highly focused output. The value of the pixel of each image is taken and compared to each other. The greatest pixel value is assigned to the corresponding pixel. This is repeated for all pixel values for sub-band multi scale coefficients and mean has been taken for low frequency multi scale coefficients.

Third Method: Absolute maximum of low frequency co-efficient Mean of sub-band orientation co-efficient

In this method we take the maximum value of the absolute of the coefficients of the pixels. This method approach has been proven the best in all above parameters. This is repeated for all pixel values for sub-band multi scale coefficients and mean has been taken for low frequency multi scale coefficients.

## 4. RESULTS

We use MATLAB software to verify the algorithm. MATLAB is a high-level language and interactive environment for numerical computation, visualization, and programming. A sample of images to be fused has been given in figure below

**Average of low frequency co-efficient and Mean of sub-band orientation co-efficient**



**Figure 4: Fused image by average of low frequency co-efficient and Mean of sub-band orientation co-efficient**

**Average of low frequency co-efficient and Min of sub-band orientation co-efficient**




**Figure 5 : Fused image by average of low frequency co-efficient and Min of sub-band orientation co-efficient**

**Average of low frequency co-efficient and Abs max of sub-band orientation co-efficient**



**Figure 6 : Fused image by average of low frequency co-efficient and abs max of sub-band orientation co-efficient**

**Table 1: Parameters values using different methods to get fused image**

Fusion method used					
	M1	M2	QABF	Entropy	STD
I <sup>st</sup> Method	6.83	6.82	0.9873	7.266	45.338
II <sup>nd</sup> Method	6.22	6.25	0.9876	7.256	45.193
III <sup>rd</sup> Method	6.11	6.05	0.9982	7.309	46.066

In the quantitative performance evaluation we evaluate fusion on the basis of statistical parameters of fused image. Several parameters can be used for evaluating the performance of fusion algorithm. In the proposed work we have used three performance evaluation metrics namely QABF, information entropy (Q) and standard deviation (s) of the original image and the fused image. Standard deviation is usually used to represent the deviation degree of the estimation and the average of the random variable. The standard deviation mainly reflects the discrete degree between the pixel gray and the mean value. The bigger the standard deviation is, the more discrete will be the distribution of gray levels. It can be calculated as

$$\sigma = \sqrt{\sum_{i=0}^{L-1} (i - \bar{i})^2 h_F(i)}, \quad i = \sum_{i=0}^{L-1} i h_F(i)$$

where  $i$ ,  $F$   $h$  and  $L$  are the grey-level index, the normalized histogram of the fused image, and the number of bins in histogram, respectively. It has been found that third method gives better result in terms of standard deviation. Entropy is one of the quantitative measures in digital image processing. Claude Shannon introduced the entropy concept in quantification of information content of the messages. Any digital image consists of the pixels arranged in several rows and many columns. Every pixel is defined by its position and also by its grey scale levels. For an image which is having  $L$  grey levels, the entropy is defined as. It is calculated as follows:

$$Q = -\sum_{i=0}^{L-1} P_i \log_2 P_i$$

where  $L$  is the number of gray level and  $P_i$  is the ratio between the number of pixels with gray values  $i$  and total number of pixels. For images with high information content the entropy will be large. In this third method proves better than the others. Similarly mutual information is a basic concept of information theory measuring the amount of information that one image contains about another. After this we used QABF parameter which describes the information about edge content. This parameter was proposed by Xydeas and Petrovic [11]. The metric measures the amount of edge information from source images to fused image. It is defined as

$$Q_p^{F/AB} = \frac{\sum_{i=1}^M \sum_{j=1}^N (Q^{AF}(i, j) W^A(i, j) + Q^{BF}(i, j) W^B(i, j))}{\sum_{i=1}^M \sum_{j=1}^N (W^A(i, j) + W^B(i, j))}$$

where,  $Q^{AF}(i, j)$  indicates the similarity between two images  $A$  and  $F$  in terms of width and direction of edge, and set

$Q^{AF}(i, j) = Q_g^{AF}(i, j) \cdot Q_a^{AF}(i, j)$ .  $Q_g^{AF}(i, j)$  and  $Q_a^{AF}(i, j)$  measure the similarity of the two images  $A$  and  $F$  in terms of strength and orientation.  $w^A(i, j)$  and  $w^B(i, j)$  are the weights of  $Q^{AF}(i, j)$  and  $Q^{BF}(i, j)$ , respectively.

$Q^{BF}(i, j)$  is similar to  $Q^{AF}(i, j)$  in terms of definition and calculation. So as conclusion it has been found that proposed methods gives good results for fusing multi-focus images and absolute value fusion rule can be preferred for fusion proces

## 5. CONCLUSION

In previous years, many kinds of multi-scale transforms have been proposed and adopted for image fusion such as pyramid decomposition, discrete wavelet transform (DWT), dual-tree complex wavelet transform (DTCWT), and discrete cosine harmonic wavelet transform (DCHWT). These transforms are widely used as wavelets become dominant filters in most of the techniques but they still have drawbacks. Wavelets are the lack of translation invariance, especially in two-dimensional (2D) signals and the poor selectivity in orientation. This can be overcome by steerable pyramid transform as one can choose the orientation before applying the filters. We have applied this method to achieve multi-focus image fusion of images which have different focus areas while capturing them. Experimental results shows that the proposed method affectively carried out fusion process as performance of the technique has been evaluated by various parameters namely mutual information, QABF factor used for edge perservance measuring and entropy etc.

## 6. REFERENCES

- [1] T. Wan, C. Zhu, Z. Qin; "Multi-focus image fusion based on robust principal component analysis," Pattern Recogn. Lett. 34 (9) (2013) 1001–1008.
- [2] A. Saha, G. Bhatnagar, Q.M.J. Wu, "Mutual spectral residual approach for multi-focus image fusion," Digit. Signal Process. 23 (4) (2013) 1121–1135.
- [3] I. De, B. Chanda, Multi-focus image fusion using a morphology-based focus measure in a quad-tree structure, Inf. Fusion 14 (2) (2013) 136–146.

- [4] S. Li, B. Yang, Multifocus image fusion using region segmentation and spatial frequency, *Image Vis. Comput.* 26 (7) (2008) 971–979.
- [5] L. Chen, J. Li, C.L.P. Chen, Regional multifocus image fusion using sparse representation, *Opt. Express* 21 (4) (2013) 5182–5197.
- [6] I. De, B. Chanda, Multi-focus image fusion using a morphology-based focus measure in a quad-tree structure, *Inf. Fusion* 14 (2) (2013) 136–146.
- [7] RichaSrivastava, Om Prakash, AshishKhare, “Biorthogonal wavelet transform based image fusion using absolute maximum fusion rule” *Information & Communication Technologies (ICT)*, 2013 IEEE Conference April 2013 Page(s): 577 – 582.
- [8] Guofeng Shao; Lixin Liu,” An Effective Wavelet-based Scheme for Multi-focus Image Fusion” Published in *Mechatronics and Automation (ICMA)*,” 2013 IEEE International Conference Aug. 2013 Page(s): 1720 – 1725.
- [9] Xu Cao; Huaxun Zhang, “A Way of Image Fusion Based on Wavelet Transform” Published in *Mobile Ad-hoc and Sensor Networks (MSN)*, 2013 IEEE Ninth International Conference on Date: 11-13 Dec. 2013 Page(s): 498 – 501.
- [10] Xiuqin Su; Fan Xu,” An Enhanced Infrared and visible Image Fusion Method Based on Wavelet Transform” *Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 2013 International Conference held on Date: 26-27 Aug. 2013 Page(s): 453 – 456.
- [11] C. Xydeas, V. Petrovic, Objective image fusion performance measure, *Electron.Lett.* 36 (4) (2000) 308–309.