



Feature-Level Multi-Focus Image Fusion Using Neural Network and Image Enhancement

By Smt.G. Mamatha, Shaik Abdul Rahim & Cyril Prasanna Raj

J.N.T University-Anantapur-India

Abstract - Image Processing applications have grown vastly in real world. Commonly due to limited depth of optical field lenses, it becomes inconceivable to obtain an image where all the objects are in focus. Image fusion deals with creating an image where all the objects are in focus. After image fusion, it plays an important role to perform other tasks of image processing such as image enhancement, image segmentation, and edge detection. This paper describes an application of Neural Network (NN), a novel feature-level multi-focus image fusion technique has been implemented, which fuses multi-focus image using classification. The image is divided into blocks. The block feature vectors are fed to feed forward NN. The trained NN is then used to fuse any pair of multi-focus images. The implemented technique used in this paper is more efficient. The comparisons of the different existing approaches along with the implementing method by calculating different parameters like PSNR, RMSE.

Keywords : Multi-focus image fusion, feed forward neural network, image Enhancement.

GJCST-E Classification: I.4.0, I.4.3



Strictly as per the compliance and regulations of:



Feature-Level Multi-Focus Image Fusion Using Neural Network and Image Enhancement

Smt.G. Mamatha^α, Shaik Abdul Rahim^σ & Cyril Prasanna Raj^ρ

Abstract - Image Processing applications have grown vastly in real world. Commonly due to limited depth of optical field lenses, it becomes inconceivable to obtain an image where all the objects are in focus. Image fusion deals with creating an image where all the objects are in focus. After image fusion, it plays an important role to perform other tasks of image processing such as image enhancement, image segmentation, and edge detection. This paper describes an application of Neural Network (NN), a novel feature-level multi-focus image fusion technique has been implemented, which fuses multi-focus image using classification. The image is divided into blocks. The block feature vectors are fed to feed forward NN. The trained NN is then used to fuse any pair of multi-focus images. The implemented technique used in this paper is more efficient. The comparisons of the different existing approaches along with the implementing method by calculating different parameters like PSNR, RMSE.

Keywords : Multi-focus image fusion, feed forward neural network, image Enhancement.

I. INTRODUCTION

Image fusion utilises information obtained from a multi-focus images of the same scene. Image processing is one of form of signal processing for which the input is an image and the output of image processing may be either an image or a set of characteristics related to the image. For the most of the image processing techniques, images of two dimensional signals is treated as input and standard image processing techniques are applied to it. The process of image fusion is performed for multi-focus and multi-sensor images of the same scene. In multi-focus images, the physical objects in the scene which are closer to the camera are in focus and the far physical object gets blurred. Adverse to it, when the far physical objects are focussed then the closer objects get blurred in the image. A hierarchical idea of image fusion has been implemented for combining significant information from multiple images into single image. The process of image fusion can be accomplished either in transformed domain or spatial domain. In spatial domain operations are performed on the pixel values. In transformed domain the images are first transformed into multiple levels of resolutions.

Author α : Assistant Professor J.N.T.U Anantapur, Department of ECE. E-mail : mamathasashi@gmail.com

Author σ : ECE J.N.T.U Anantapur.

E-mail : rahimengg1001@gmail.com

Author ρ : Professor Department of EEE M.S.R.S.A.S Bangalore India.

E-mail : cyrilyahoo@gmail.com

Information fusion can be performed at any level of the image information representation corresponding to other forms of information fusion, image fusion is usually performed at one of the three different processing levels they are Pixel, Feature and Decision Level [5]. The pixel level image fusion is also known as signal level image fusion which represents fusion at the lowest processing level, that is operations such as maximum or mean(average) are applied to the pixel values of the source images to generate the fused image. Feature level image fusion is also known as object level image fusion where fused features and object labels and information that have already extracted from individual input images. Decision level is also known as symbol level, the objects in the input image are first detected and then the suitable fusion algorithm the fused image is generated. In the field of Image Processing, image fusion has received a significant importance for medical imaging, military applications, forensic, remote sensing.

A number of image fusion techniques have been exhibited in the literature. In addition of simple pixel level image fusion techniques. We find the complex techniques such as Laplacian Pyramid [2], Morphological pyramid [6], fusion based on PCA [3], Discrete wavelet Transform (DWT) [1]. These fusion techniques have different advantages and disadvantages such as liner wavelets during image decomposition the fused image doesn't preserve the original data. Likewise due to low-pass filtering of wavelets, the edges in the image becomes smooth and hence the contrast in fused image is decreased.

In this paper, we have implemented a method for multi-focus image fusion. The implemented method is discussed in section II. In section III, the quantitative measures used to evaluate the performance of the implemented method are described. Section IV covers the experiments details and section V concludes the study.

II. IMPLEMENTING METHOD

Different images are acquired from the Image Processing websites. From the acquired images consider one image, for that image generate two source images from original image that is one is left focused and right blurred other one right focused and left blurred. Every image is divided into blocks. The block size plays a significant role in differentiating the blurred

and un-blurred regions from each other. After dividing the image into blocks, the feature values of block of all the images are calculated and feature file is generated. A comfortable number of feature vectors are used to train NN. The trained NN is then used to fuse any set of multi-focus images. Image data set, feature selection and implemented algorithm are discussed in the following sections.

a) *Creating Image Dataset*

In the implementing method, we created an image-data set of ten grayscale images. These images are acquired from the different image processing websites. For each image in the data set, we generated its two versions of the same size. In the first versions, the left half of the image is blurred and right image is focused. A similar process is performed in the right image is blurred and left image is focused. The blurred versions are generated by Gaussian blurring of radius 1.5. In implementing method experimentation, we resize all the images into 256*256 resolutions.

b) *Feature Extraction*

In feature-level image fusion, the selection of different features is an important task. The blurred objects in an image reduce its clearness. In multi-focus images, some objects are in focus and some objects are blurred. In this paper, we extract five features from each image block to represent its clearness. These are the Variance, Energy of gradient, Contrast visibility, Spatial frequency and canny edge information. From the figure (1), we calculate the blurriness of Gaussian radius for source image, Left blur and Right focused, and Left focused and Right blur. The value of features in the image against blurriness is given in table (I). If the blurriness is increased the values of energy gradient, spatial frequency and edge information are reduced.

1) *Contrast Visibility* : It calculates the deviation of a block of pixels from the block's mean value. Therefore it relates to the clearness level of the block. The visibility of the image block is obtained using equation (1)

$$CV = \frac{1}{p * q} \sum_{(i,j) \in Bk} \left| \frac{X(i, j) - \mu_k}{\mu_k} \right| \quad (1)$$

Here μ_k and $p * q$ are the mean and size of the block B_k respectively.

2) *Spatial Frequency*: Spatial frequency measure the activity level in an image, it used to calculate the frequency changes along rows and columns of the image. Spatial frequency is measured using equation (2).

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

Where

$$RF = \sqrt{\frac{1}{p * q} \sum_{i=1}^p \sum_{j=2}^q [X(i, j) - X(i, j-1)]^2} \quad \text{and}$$

$$CF = \sqrt{\frac{1}{p * q} \sum_{i=1}^p \sum_{j=2}^q [X(i, j) - X(i-1, j)]^2}$$

Here X is the image and $p * q$ is the image size. A large value of spatial frequency describes the large information level in the image and therefore it measures the clearness of the image.

3) *Variance*: Variance is used to measure the extent of focus in an image block. It is calculated using equation (3)

$$Variance = \frac{1}{p * q} \sum_{i=1}^p \sum_{j=1}^q (X(i, j) - \mu)^2 \quad (3)$$

Here μ is the mean value of the block image and $p * q$ is the image size. A high value of variance shows the greater extent of focus in the image block.

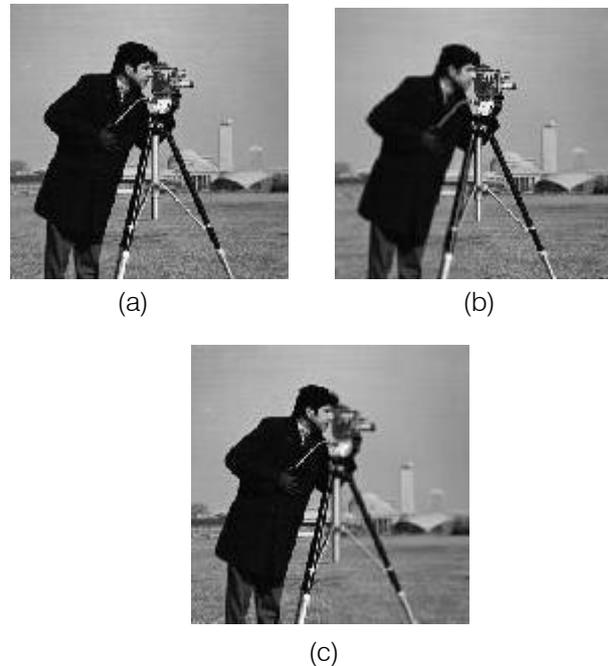


Figure 1: Cameraman image (a) Original Image (b) LBRF (c) LFRB

Table 1: Feature values for Original Image, LBRF and LFRB

Figure	Variance	EG	SPF	CV	Edge
1 (a)	148.41	26274	0.0286	0.1057	5747
1 (b)	146.18	21752	0.0264	0.1003	5402
1 (c)	145.19	17854	0.0243	0.0992	5106

4) Energy of Gradient (EG): It is also used to measure the amount of focus in an. It is calculated using equation (4).

$$EG = \sum_{i=1}^{p-1} \sum_{j=1}^{q-1} (r_i^2 + r_j^2) \quad (4)$$

Where $r_i = r(i+1, j) - r(i, j)$

and $r_j = r(i, j+1) - r(i, j)$

Here p and q represent the dimensions of the image block. A high value of energy of gradient shows greater amount of focus in the image block.

5) Edge Information: The edge pixels can be found in the image block by using canny edge detector. It returns 1 if the current pixel belongs to some edge in the image otherwise it returns 0. The edge feature is just the number of edge pixels contained within the image block.

c) Artificial Neural Networks

Many Neural Network models have been implemented for tackling a diverse range of problems*, including pattern classification. The fusion we examine here can be considered as classification problem. Here we have considered a NN applications model, namely the PNN (Probabilistic Neural Network). The basic idea underlying NN is to overlap localized receptive fields of the hidden units to create arbitrarily complex nonlinear ties. The normal architecture consists of one hidden layer and one output layer. Each hidden unit corresponding to a kernel or basis function of the input vector x, and is usually of the Gaussian form. The basic architecture of feed forward NN is shown below

$$Z(x) = \exp(-\|x - c\|^2 / \sigma^2)$$

Here, c is the position of the hidden unit and is a user-defined width that controls its spread. For PNN a hidden unit is positioned at every training data point.

d) Neural Network Algorithm

The algorithm first decomposes the source images into blocks. Given two of these blocks (one from each source image), a neural network is trained to determine which one is clearer. Fusion then proceeds by selecting the clearer block in constructing the final image. The fusion result of DWT is shift dependent. The use of image blocks on the other hand, avoids this problem even if there is object movement or mis-registration in the source images, each object will still be in better focus in one of the source images. In detail, stepwise working of the implemented method is given under. 1) LFi is the left-focused and RFi is the right-focused versions of the ith image in the dataset in section(II-A). 2) Divide the versions LFi and RFi of every image in the dataset into k number blocks of the size

M*N. 3) Create the features file for all LFi and RFi according to the features discussed in section (II-B). Here j=1,2,3...k. For all i, there are two sets of features values for every block j named as FSLFi and FSRFi each of which contains five feature values. Subtract the features values of block j of RFi and include this pattern in feature file. Normalise the feature value between [0 1]. 4) Assign the class value to every block j of ith image. If block j is visible in LFi then assign it class value 1 otherwise give it a class value -1. In case of class value -1, block j is visible in RFi. 5) Train a neural network to determine whether LFi or RFi is clearer. Identify the clearness of all the blocks of any pair multi-focus images to be fused. 6) Fuse the given pair of multi-focus images block by block according to the classification results of the neural network. Such that Output of NN for block J If>0,select J from left-focused Image If<0,select j from right-focused Image The block diagram of the implemented method is shown in figure (2).

III. QUANTITATIVE MEASURES

There are different quantitative measures which are used to evaluate the performance of the fusion techniques. These are PSNR (Peak Signal to noise ratio), RMSE (root mean square error), Entropy, Correlation Coefficient, MAE (mean absolute error).

a) PSNR

Determines the degree of resemblance between reference images fused image and fused image. A bigger value shows good fusion results.

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

L denotes to number of gray level in the image.

b) RMSE

Calculate the deviation between the pixel values of reference image and fused image. A lesser value shows the good fusion results.

$$RMSE = \sqrt{\frac{1}{p \cdot q} \sum_{i=1}^p \sum_{j=1}^q [R(i, j) - F(i, j)]^2}$$

Here R, F are the reference and fused images respectively.p*q is the image size.

c) Entropy

Quantifies the quantity of information contained in the fused image. A bigger value shows good fusion results.

$$H = - \sum_{i=0}^{L-1} h_F(i) \log_2 h_F(i)$$

Here h_F is the normalized histogram of fused image and L is the number of gray levels.

d) *Correlation Coefficient*

The correlation coefficient matrix represents the normalized measure of the strength of linear relationship between variables.

$$CC = \frac{\sum_{t=1}^N (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^N (x_t - \bar{x})^2}$$

Where x_t is a data value at time step t, k is the lag.

e) *MAE*

It is used to calculate the mean absolute error between reference image and fused image.

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

Where $F(i, j)$ the predicted is fused image and $R(i, j)$ is the true value fused image.

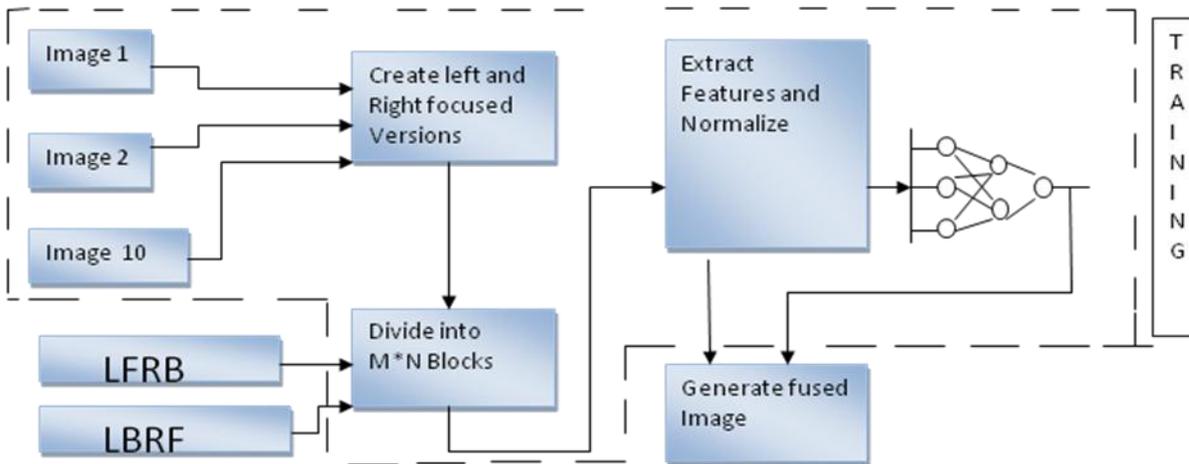


Figure 2: Block Diagram of the Implementing method

IV. EXPERIMENTS AND RESULTS

Image fusion is performed. The implemented technique used in this paper is more efficient and useful, to highlight the efficiency. We have performed broad experimentation on this technique. We trained the feed forward neural network with different number of hidden layers and with different number of neurons on each layer. The results of the implementing technique are compared with different existing methods including DWT, PCA, Laplacian pyramid, Morphological processing based image fusion techniques. To calculate the performance of the implementing technique, the results for two different pairs of multi-focus images are obtained including Pepsi, balloon, and cameraman images.

a) *Difference between DWT based Image fusion and the Implementing Technique*

Gonzalo pajares, implemented a DWT based image fusion technique to perform multi-focus image fusion. The fusion result Discrete Wavelet Transform is shiftdependent. The use of image block, on the other hand avoids the problem of shift dependent. We have used five different features like (SPF, EG, CV, Variance, Edge) to calculate the clearness of a block more accurately as compared to DWT. Even though if there is

object misregistration or movement in the input images, each object will still be in finer focus in one of the input images. Thus, in the fused result, all the blocks covering a particular object will come from the same input image and its clarity will not be affected due to any misregistration problem.

b) *Visual Comparison Assessments*

The visible comparisons has shown in below figures (3) for Pepsi images, figure (4) for balloon images, figure (5) for cameraman images. All the images are of size 256*256.



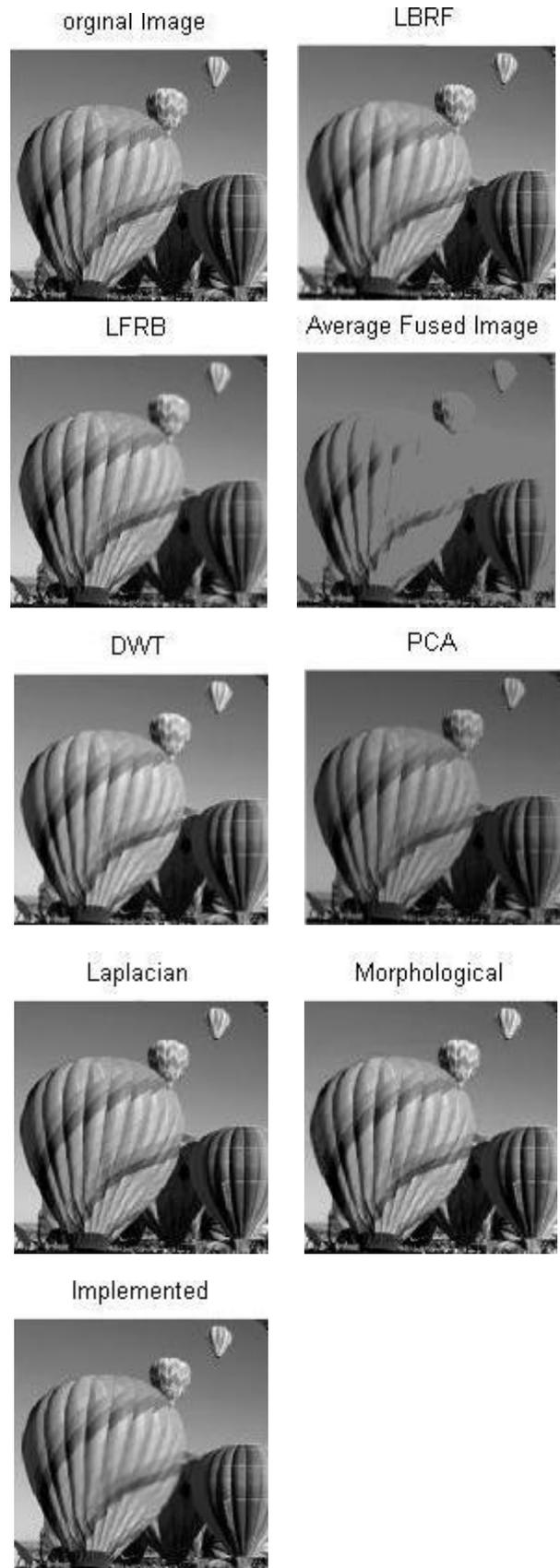


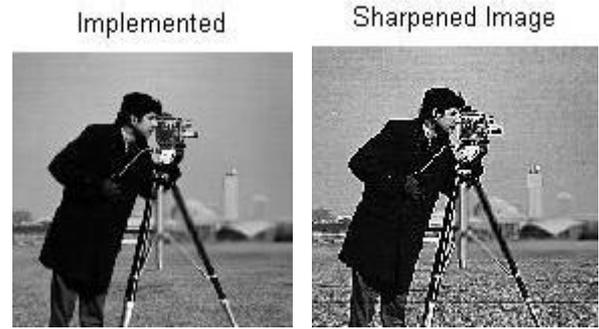
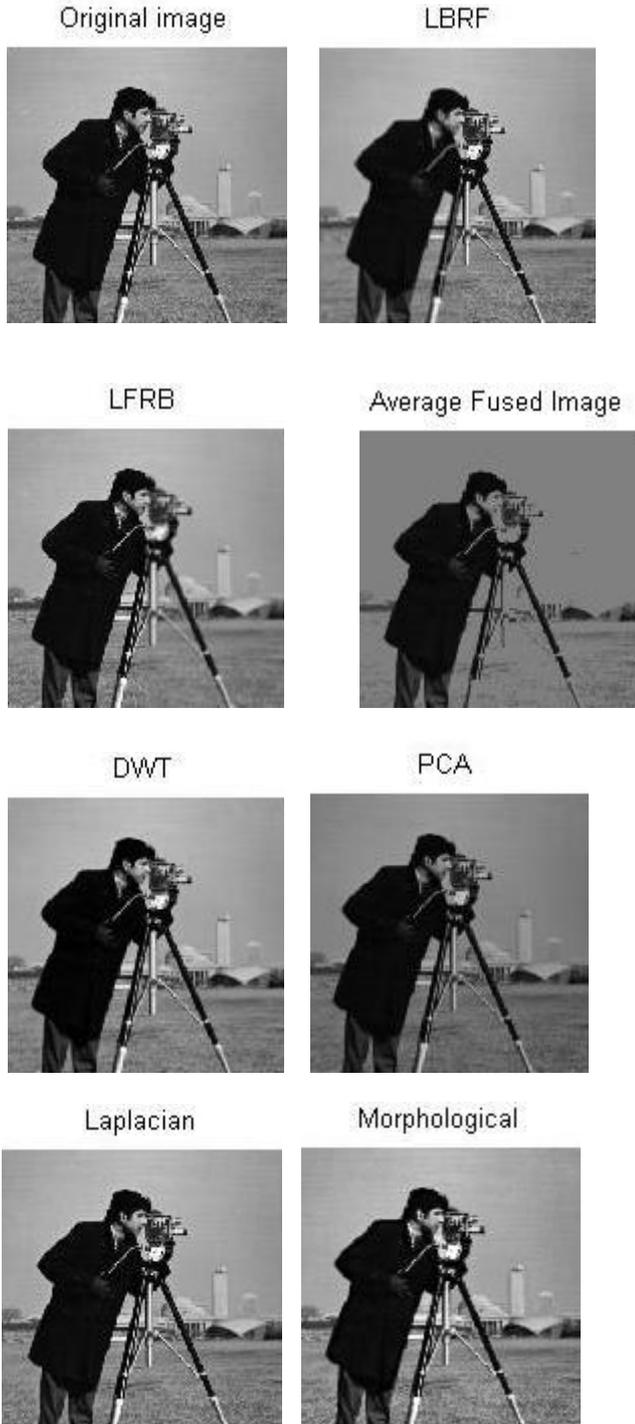
Table 2: Results of Quantitative Measures for Pepsi Image

Method	PSNR	RMSE	Entro	Corre	MAE
Average	17.64	33.42	4.4408	0.955	20.57
DWT	5.621	133.0	0.9978	0.998	118.7
PCA	19.52	9.73	6.9321	0.998	117.1
Laplacian	5.88	129.5	7.2997	0.996	117.0
Morphological	5.88	129.5	7.2815	0.991	117.0
Implementing	34.80	4.635	7.2913	0.9966	0.0104

Method	PSNR	RMSE	Entro	Corre	MAE
Average	19.79	26.124	5.0694	0.9182	12.62
DWT	6.322	123.14	0.975	0.9946	111.55
PCA	20.07	10.12	7.214	0.9946	111.20
Laplacian	6.356	122.66	7.598	0.9892	111.11
Morphological	6.357	122.65	7.602	0.9850	111.10
Implementing	35.89	4.0886	7.5827	0.9971	0.0947

Table 3: Results of Quantitative Measures for Ballon Image

PSNR result for Cameraman images:



PSNR=31.26

High performance.

Texture Calculations:

Texture Features	Source Image	Fused Image Implemented
Energy	1.1785e+009	1.1726e+009
Dissimilarity	688002945	687848288

V. CONCLUSION

In this implemented technique, a feature level focus image fusion has been implemented in this paper. In this method we have trained the feed forward neural network with the block features of pairs of multi-focus images. A feature set including SF, CV, edges, variance, and EG is used to define clarity of the image block. The trained neural network was then used to fuse any pair of multi-focus images. Experimentation results show that the implemented technique performs better than the existing techniques. The fusion result of Discrete Wavelet Transform is shift dependent. The use of image block, on the other hand avoids the problem of shift dependent.

REFERENCES RÉFÉRENCES REFERENCIAS

1. H.Li, S.Manjunath and S.K.Mitra, "Multi-sensor image fusion using the wavelet transform" in Graphical Models and Image processing, vol57, no.3, pp.235-245, 1995.
2. A.Toet, "Image fusion by a ratio of low pass pyramid" in pattern Recognition Letters, vol.9, no.4,pp.245-253, 1989.
3. V.P.S.Naidu and J.R.Raol, "Pixel-level Image Fusion using Wavelets and Principal Component Analysis" in Defence Science Journal, vol.58, no.3,pp.338-352,May 2008.
4. Shutao Li,James J.Knok, "Multi-focus image fusion using artificial neural networks" in pattern Recognition letters,vol 23,pp 985-997.
5. Gonzalo Pajares and Jesus Manuel de la Cruz, "A wavelet-based Image Fusion Tutorial" in pattern Recognition, vol 37, no.9, pp.1855-1872, 2004.
6. Yufeng Zheng, Edward A. Essock and Bruce C.Hansen, "An Advanced Image Fusion Algorithm Based on Wavelet Transform-Incorporation with

- PCA and morphological processing” in Processing of the SPIE,vol 5298,pp.177-187,2004.
- (2002) The IEEE website. [Online]. Available: <http://www.ieee.org/>. [<http://www.doc.ic.ac.uk/~nd/surprise96/journal/vol14/cs11/reporthtml#Introductiontoneuralnetworks>].



This page is intentionally left blank

