International Journal of the Physical Sciences Vol. 6(20), pp. 4838-4847, 23 September, 2011
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS11.731
ISSN 1992 - 1950 ©2011 Academic Journals

Full Length Research Paper

Feature classification for multi-focus image fusion

Abdul Basit Siddiqui[1], Muhammad Rashid[1]*, M. Arfan Jaffar[2], Ayyaz Hussain[3] and Anwar M. Mirza[2]

[1]Foundation University Institute of Engineering and Management Science, New Lalazar, Rawalpindi, Pakistan.
[2]National University of Computer and Emerging Sciences, Islamabad, Pakistan.
[3]International Islamic University, Islamabad, Pakistan.

Accepted 29 July, 2011

Image processing techniques have witnessed increased usage in various real world applications. For any image processing technique, such as image segmentation, restoration, edge detection, stereo matching etc., to be applied successfully, the image under consideration must contain all of the scene objects in focus. Usually, due to inadequate depth of field of optical lenses, especially with larger focal length, it becomes impossible to obtain an image in which all of the objects are in focus. Image fusion deals with creating an image by combining portions from other images to obtain an image in which all of the objects are in focus. In this paper, a novel feature-level multi-focus image fusion technique has been proposed which fuses multi-focus images using classification. Ten pairs of multi-focus images are first divided into blocks. The optimal block size for every image is found adaptively. The block feature vectors are fed to feed forward neural network. The trained neural network is then used to fuse any pair of multi-focus images. The results of extensive experimentation performed are presented to highlight the efficiency and usefulness of the proposed technique.

Key words: Multi-focus image fusion, feed forward neural network, feature classification, genetic algorithm.

INTRODUCTION

Image fusion is a sub-field of image processing in which portions of more than one images are fused together creating an image where all the objects are in focus. Image fusion is of significant importance due to its application in medical science, forensic and defense departments. The process of image fusion is performed for multi-sensor and multi-focus images of the same scene. Multi-sensor images of the same scene are captured by different sensors whereas multi-focus images are captured by the same sensor. In multi-focus images, the objects in the scene which are closer to the camera are in focus and the farther objects get blurred. Contrary to it, when the farther objects are focused then closer objects get blurred in the image. To achieve an image where all the objects are in focus, the process of images fusion is performed either in spatial domain or in transformed domain. Spatial domain includes the techniques which directly incorporate the pixel values. In transformed domain, the images are first transformed into multiple levels of resolutions. An image often contains physically relevant features at many different scales or resolutions. Multi-scale or multi-resolution approaches provide a means to exploit this fact (De and Chanda, 2006). After applying certain operations on the transformed images, the fused image is created by taking the inverse transform.

Image fusion is generally performed at three different levels of information representation including pixel level, feature level and decision level (Pajares and de la Cruz, 2004). In pixel-level image fusion, simple mathematical operations such as max (maximum) or mean (average) are applied on the pixel values of the source images to generate fused image. However, these techniques usually smooth the sharp edges or leave the blurring effects in the fused image. In feature level multi-focus image fusion, the source images are first segmented into different regions and then the feature values of these regions are calculated. Using some fusion rule, the regions

*Corresponding author. E-mail: rashid.nuces@gmail.com. Tel: +923335137267.

Abbreviations: DWT, Discrete wavelet; **GA,** genetic algorithm; **EOG,** Energy of gradient; **SF**, spatial frequency; **SD,** standard deviation; **MI,** mutual information; **PSNR,** peak signal to noise ration; **RMSE,** root mean square error.

are selected to generate the fused image. In decision level image fusion, the objects in the source images are first detected and then by using some suitable fusion algorithm, the fused image is generated.

A number of image fusion techniques have been presented in literature. In addition to the simple pixel level image fusion techniques, we find the complex techniques such as Laplacian Pyramid (Toet, 1989), fusion based on PCA (Naidu and Raol, 2008), discrete wavelet (DWT) based image fusion (Li et al., 1995), Neural Network based image fusion (Li et al., 2002) and advance DWT-based image fusion (Zheng et al., 2004). These techniques have different merits and demerits such as linear wavelets like Haar wavelet during the image decomposition does not preserve the original data (Heijmans and Goutsias, 2000). Similarly, due to low-pass filtering process of wavelets, the edges in the image become smooth and hence, the contrast in the fused image is decreased.

In this paper, we have proposed a new method for multi-focus image fusion which incorporates Genetic algorithm (GA) to find optimal block size.

GENETIC ALGORITHM (GA)

GA is based on Darwin's theory of evolution. The idea of evolutionary computing was introduced by Rechenberg in 1960s. In GA, initially a population of chromosomes (solutions to the problem) is initialized within the search space. Using the previous population, new population is created which is comprised of better solutions. The selection of the solutions for new population is made on the basis of their fitness values. Solutions of greater fitness have more chances to reproduce the new solutions. Genetic operator crossover and mutation with a certain probability are used on the parent solutions to generate off-springs. This iterative procedure is carried out until some stopping criterion is reached. The stopping criteria include the number of populations or maximum fitness achieved. An introduction about the working of GA can be found in Goldberg (1989).

PROPOSED METHOD

We used an image set of 10 different images to train our neural network. Every image is first divided into number of blocks. The block size plays an important role in distinguishing the blurred and un-blurred regions from each other. To accomplish this task, we run the algorithm for finding adaptive block size using GA introduced in Zhang et al. (2005) for every image in the image set. Zhang et al. (2005) used a population of chromosomes where every chromosome represents the width and length of the block. After dividing the images into blocks, the feature values of every block of all the images are calculated and a features file is created. A sufficient number of feature vectors are used to train the neural network. The trained neural network is then used to fuse any set of multi-focus images.

Creating image dataset

In the proposed method, we first created an image-set of 10 grayscale images. This image-set is shown in the Figure 1. These images are taken from different image processing websites. The images are of different scenes and backgrounds. For every image in the set, two versions of the same size were created. In the first version, some of the regions are randomly selected in the left half of the image and are blurred. A similar process is performed in the right half of the image in the second version. The blurred versions are generated by Gaussian blurring of radius 1.5. For the image set of 10 grayscale images, we created 20 versions of blurred images. For proposed method experimentation, we resized all the images into 480x640 resolutions.

Features selection

In feature-level image fusion, the selection of different features is an important task. In multi-focus images, some of the objects are clear (in focus) and some objects are blurred (out of focus). The blurred objects in an image reduce its clearness. We have used five different features to characterize the information level contained in a specific portion of the image. This features set includes variance, energy of gradient, contrast visibility, spatial frequency and canny edge information. Figure 2 shows how the blurriness of increasing Gaussian radius affects the clearness of the image. The values of the features in the image against blurriness of different degrees are given in Table 1.

We observe from Figure 1 that with the increasing blurriness, the clearness of the image is reduced and the identification of different objects in the image becomes difficult.

Features values given in Table 1 showed the degradation of the original image. The values of energy of gradient, spatial frequency and edge information are significantly reduced.

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.

$$VI = \frac{1}{m \times n} \sum_{(i,j) \in B_k} \frac{|I(i,j) - \mu_k|}{\mu_k} \quad (1)$$

Where μ_k and $m \times n$ are the mean and size of the block B_k, respectively.

Spatial frequency

Spatial frequency measures the activity level in an image. It is used to calculate the frequency changes along rows and columns of the image. Spatial frequency is measured using equation (2).

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

Where,

$$RF = \sqrt{\frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=2}^{n} \left[I(i,j) - I(i,j-1) \right]^2}$$

Figure 1. Image set used to train neural network.

Table 1. Feature values with increasing blurriness in BARM image shown in Figure 1.

Figure	Variance	EOG	SF	VI	Edge
a	195.74	463.37	21.57	0.326	2422
b	189.09	376.72	17.925	0.318	2352
c	176.98	205.61	14.346	0.308	2225
D	170.44	192.27	13.851	0.3067	2162

$$CF = \sqrt{\frac{1}{m \times n} \sum_{j=1}^{n} \sum_{i=2}^{m} \left[I(i, j) - I(i-1, j) \right]^2}$$

Here, I is the image and $m \times n$ 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.

Variance

Variance is used to measure the extent of focus in an image block. It is calculated using Equation 3:

$$Variance = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} (I(i, j) - \mu)^2 \qquad 3$$

Where, μ is the mean value of the block image and mxn is the image size. A high value of variance shows the greater extent of focus in the image block.

Energy of gradient (EOG)

It is also used to measure the amount of focus in an image block. It is calculated using Equation 4.

$$EOG = \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} (f_i^{\,2} + f_j^{\,2}) \qquad 4$$

$$where \quad f_i = f(i+1, j) - f(i, j)$$

$$and \quad f_j = f(i, j+1) - f(i, j)$$

Here, m and n represent the dimensions of the image block. A high value of energy of gradient shows greater amount of focus in the image block.

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.

Proposed algorithm

Stepwise working of the proposed method is as follows:

Find the optimal block size for each set of LF_i and RF_i using (Zhang et al., 2005). LF_i is the left-focused and RF_i is the right-focused versions of the ith image in the dataset discussed in section (2.1). where $i = 1, 2, 3, \ldots, 10$.

Divide the versions LF_i and RF_i of every image in the dataset into K number of blocks of size MxN.

Create the features file for all LF_{ij} and RF_{ij} according to the features discussed in section (2.2). Here $j = 1, 2, 3, \ldots, K$. For

Figure 2. BARM image (a) Original Image (b) blurred with radius 0.5 (c) blurred with radius 1.0 and (d) blurred with radius 1.5.

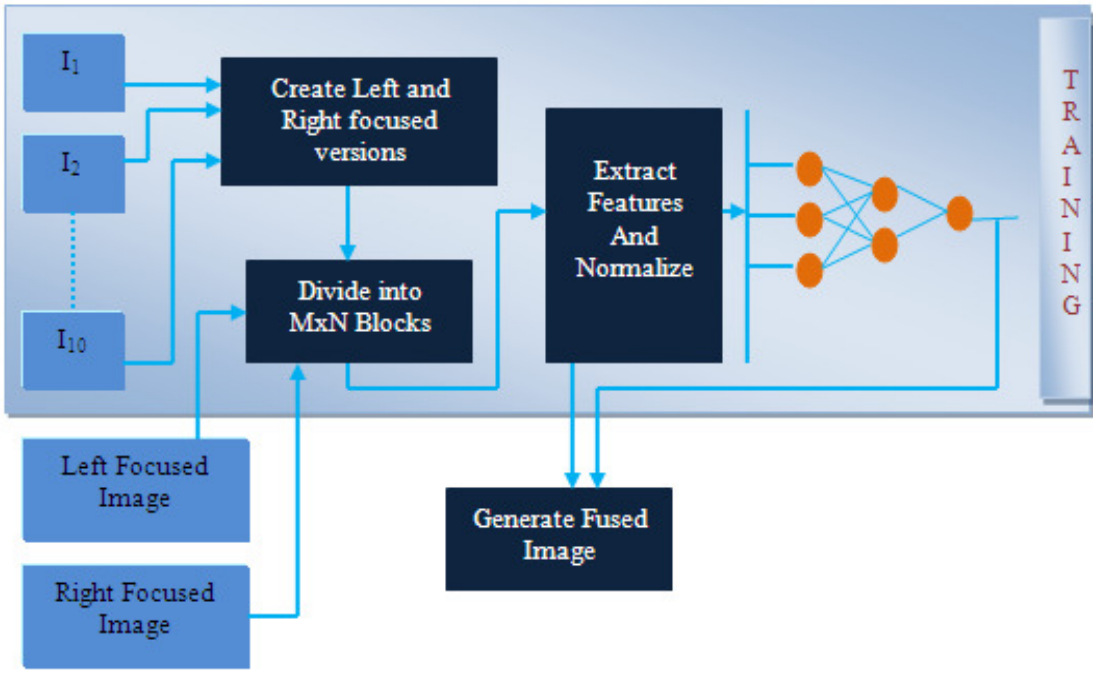

Figure 3. Block diagram of the proposed method.

all i, there are two sets of features values for every block j named as $FSLF_{ij}$ and $FSRF_{ij}$ each of which contains five feature values. Subtract the features values of block j of LF_i from the corresponding feature values of block j of RF_i and include this pattern in features file. Normalize the feature values between [0 1]. Assign the class value to every block j of ith image. If block j is visible in LF_i then assign it class value 1 otherwise give it a class value -1. In case of class value -1, block j is visible in RF_i. Create a neural network with adequate number of layers and neurons. Train the newly created neural network with adequate number of patterns selected from features file created in step 2. By using the trained neural network, identify the clearness of all the blocks of any pair of multi-focus images to be fused. Fuse the given pair of multi-focus images block by block according to the classification results of the neural network such that;

$$Output \quad of \quad NN \quad for \quad block \quad j =$$

$$\begin{cases} if \ > 0, & Select \quad j \quad from \quad left - focused \quad image \\ if \ < 0, & Select \quad j \quad from \quad right - focused \quad image \end{cases}$$

The block diagram of the proposed method is shown in Figure 3.

Quantitative measures

There are different quantitative measures which are used to evaluate the performance of the fusion techniques. We used five different measures including Root Mean Square Error (RMSE), Peak Signal to Noise Ration (PSNR), Mutual Information (MI), Correlation and Mean Absolute Error when the reference image is available. For blind image fusion (reference image is not available), we used Entropy, Mutual information (MI), Standard deviation (SD),

Spatial frequency (SF). These measures are defined in Table 2.

In addition of the performance metrics listed in Table 2, some other performance measures usually used to check performance of any fused algorithm can be found (Naidu et al., 2003; Wang and Bovik, 2002; Blum and Zheng, 2006; Cvejic et al., 2005).

EXPERIMENTS AND RESULTS

Image fusion is performed in two different situations. In first situation, the reference image is available and in second situation, the reference image is not available (blind image fusion). We exercised the feed forward neural network with different number of hidden layers and with different number of neurons on each layer. We found the best results with one hidden layer of 30 neurons. The learning rate α and the threshold for mean square error are kept as 0.01. The results of the proposed technique are compared with different existing methods including DWT, aDWT, PCA and Laplacian Pyramid based image fusion techniques. The performance of an existing Probabilistic Neural Network based technique is also compared with the results of the proposed technique. The experimentation results are obtained when the reference image is available and when it is not available (blind image fusion). In order to evaluate the performance of the proposed technique, the results for four different pairs of multi-focus images are obtained including balloon, lab, clock and Pepsi can images.

PNN-based image fusion and the proposed technique variation

Li et al. (2002) proposed a probabilistic neural network based technique to perform multi-focus image fusion. They trained a neural network to classify the selection of blocks of the two source images to generate the fused image. In their technique, the source images were divided into the blocks of fixed size, 32x32. The block size is an important factor to achieve good fusion results. Fixed size blocks may separate the blurred and un-blurred regions within one pair of multi-focus images but for some other pair of multi-focus images, the contents of the image block are partially blurred. The size of the block varies from image to image because different images have different blurred regions. In the proposed method, for every pair of multi-focus images, an optimal block size is found using the technique given in Zhang et al. (2005). We have used five different features to calculate the clearness of a block more accurately as compared to three features used by Li et al. (2002)

A major difference between the proposed method and the existing PNN-based image fusion technique is the training of the neural network. Li et al. (2002) in their technique create and train a new neural network for every pair of multi-focus images which is time consuming. In the proposed method, we trained the neural network using the block features of ten different pair of multi-focus images. Once the classifier is obtained then it can be used to fuse any pair of multi-focus images.

Quantitative assessments and the visual comparison

When the reference image is available

For balloon and lab images, the reference images are available. A visual comparison is shown in Figures 4 and 5 for balloon and lab images. Both the balloon and lab images are of size, 480x640.

We used three different quantitative measures to evaluate the performance of fusion classifier proposed in this paper. The experimentation results obtained for balloon and lab images are summarized in Table 3.

Visual examination of Figures 4 and 5 shows that, the fused images obtained by the proposed technique are clearer than the fused images obtained by using the existing techniques. The performance of the proposed technique can also be verified from the results of different quantitative measures given in Table 3.

When the reference Image is not available (Blind Image Fusion)

We have used clock and Pepsi can images in this category. The clock and Pepsi can image sizes are 256x256 and 512x512, respectively. When the reference image is not available then the performance of the fusion process is evaluated on the basis of different set of quantitative measures. Figures 6 and 7, provides a visual comparison of the proposed technique with the existing techniques.

In case of Pepsi can image, block effects are visible at the table edge in the fused image generated by PNN-based technique. In PNN-based technique, the source images are divided into parts using a block of fixed size and hence it leaves the block effects in the fused image.

Table 4 provides some statistics to evaluate the performance of the proposed technique when the reference image is not available. These statistics include the results of different quantitative measures. In case of fused image generated by Laplacian Pyramid-based technique, the value of spatial frequency is greater than the proposed method. However, in general, the proposed technique performs better than existing techniques.

CONCLUSION

In this paper, a feature-level block-based multi-focus image fusion technique is proposed. A feed forward neural network is first trained with the block features of ten pairs of multi-focus images. A feature set including spatial frequency, contrast visibility, edges, variance and energy of gradient is used to define the clarity of the

Table 2. Performance metrics used for image fusion.

Metric	Formula	Description
RMSE	$RMSE = \sqrt{\frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} [R(i,j) - F(i,j)]^2}$ Where, R and F are the reference and fused images, respectively. mxn is the image size.	Calculates the deviation between the pixel values of reference image and fused image. A lesser value shows good fusion results.
PSNR	$PSNR = 20 \log_{10} \left[\frac{L^2}{\frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} [R(i,j) - F(i,j)]^2} \right]$ L denotes to number of gray level in the image.	Determines the degree of resemblance between reference and fused image. A bigger value shows good fusion results.
MI	$MI = \sum_{i=1}^{m} \sum_{j=1}^{n} h_{R,F}(i,j) \log_2 \left[\frac{h_{R,F}(i,j)}{h_R(i,j) h_F(i,j)} \right]$ Here $h_{R,F}$, h_R, h_F are the joint, reference and fused images histograms respectively. Mutual Information can also be used when reference image is not available. In that case, it is the sum of the mutual information retrieved by fused image from image A and image B.	Determines how much information the fused image retrieved from the input source images. A bigger value shows good fusion results.
CORR	$CORR = \frac{2 C_{rf}}{C_r + C_f}$ where $C_r = \sum_{i=1}^{m} \sum_{j=1}^{n} R(i,j)^2$, $C_f = \sum_{i=1}^{m} \sum_{j=1}^{n} F(i,j)^2$ and $C_{rf} = \sum_{i=1}^{m} \sum_{j=1}^{n} R(i,j) \, F(i,j)$	It is used to determine how reference and fused images are identical to each other. Its value varies between 0 and 1. If it returns 1 then it means both reference and fused images are absolutely similar.
MAE	$MAE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j}^{N} \lvert R(i,j) - F(i,j) \rvert$ where M, N are the number of rows and columns of the image respectively.	It is used to calculate the mean absolute error between reference and fused image.
Entropy	$H = -\sum_{i=0}^{L-1} h_F(i) \log_2 h_F(i)$ where h_F is the normalized histogram of fused image and L is the number of gray levels.	Quantifies the quantity of information contained in the fused image. A bigger value shows good fusion results.

Table 2. Contd.

SD	$SD=\sqrt{\sum_{i=0}^{L}(i-i')^2\, h_F(i)} \quad where \quad i'=\sum_{i=0}^{L} i\, h_F$ Where, h_F is the normalized histogram of fused image and L is the number of gray levels.	Measures the contrast in the fused image. A well contrast image has high standard deviation.
SF	Spatial Frequency is described in section 3.2	-------

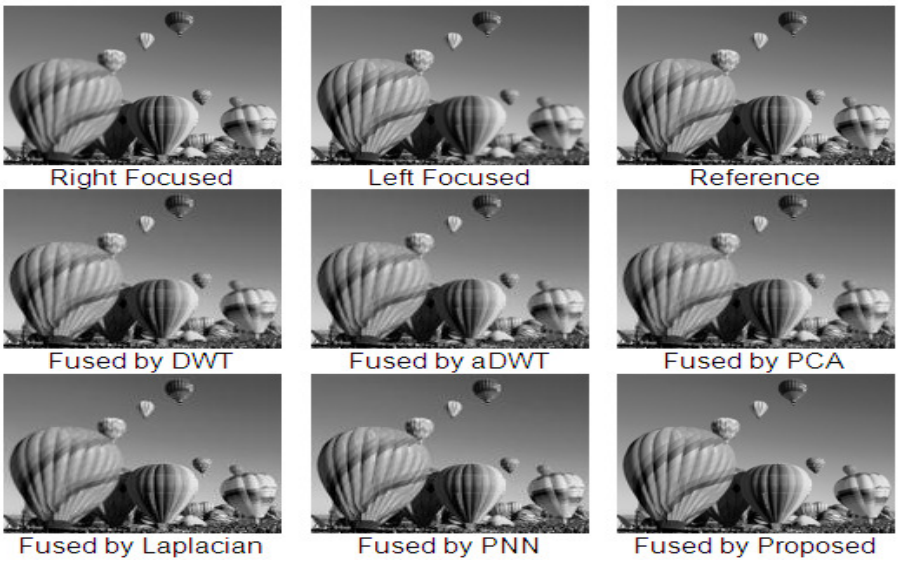

Figure 4. Balloon images fused by different image fusion techniques and the proposed method.

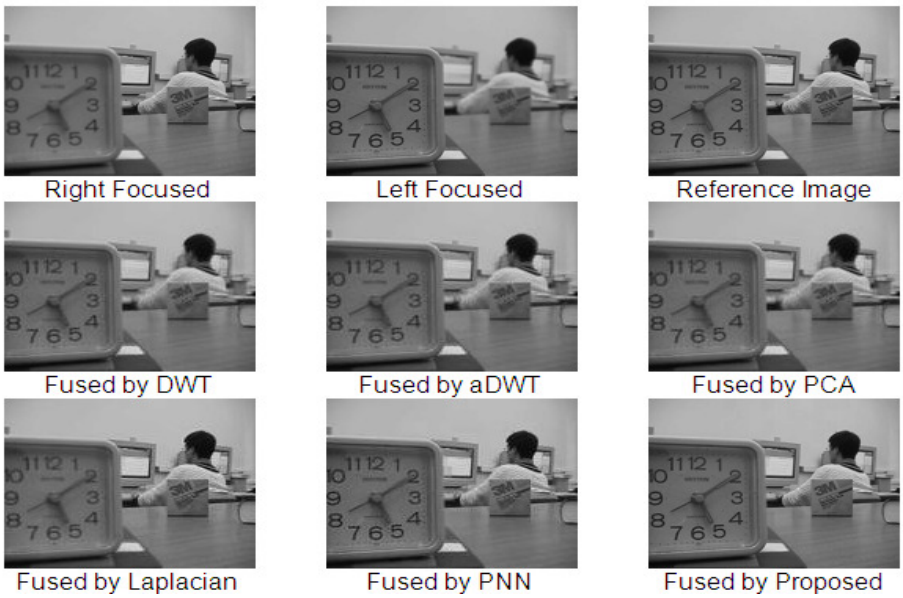

Figure 5. Lab images fused by different image fusion techniques and the proposed method.

Table 3. Experimentation results of different quantitative measures for fused balloon and lab images.

Image	Method	Quality measures				
		RMSE	PSNR	MAE	CORR	MI
Balloon	DWT	5.1025	33.9751	2.1629	0.9991	12.6004
	aDWT	5.0781	34.0168	2.1568	0.9992	12.6122
	PCA	6.0099	32.5535	2.5488	0.9988	12.0784
	Laplacian	2.8222	39.1191	1.1992	0.9997	16.2871
	PNN	0.2634	39.8945	1.0415	0.9996	18.6192
	Proposed	0.1967	62.2550	0.9981	0.9998	24.3553
Lab	DWT	6.8450	31.4234	3.5299	0.9986	6.6182
	aDWT	6.8088	31.4694	3.5168	0.9987	6.6366
	PCA	7.0536	31.1625	3.5577	0.9986	6.7882
	Laplacian	4.2743	35.5135	1.7704	0.9995	9.3071
	PNN	3.2999	37.7607	1.9317	0.9988	14.4189
	Proposed	2.1985	41.2883	1.4221	0.9997	15.2816

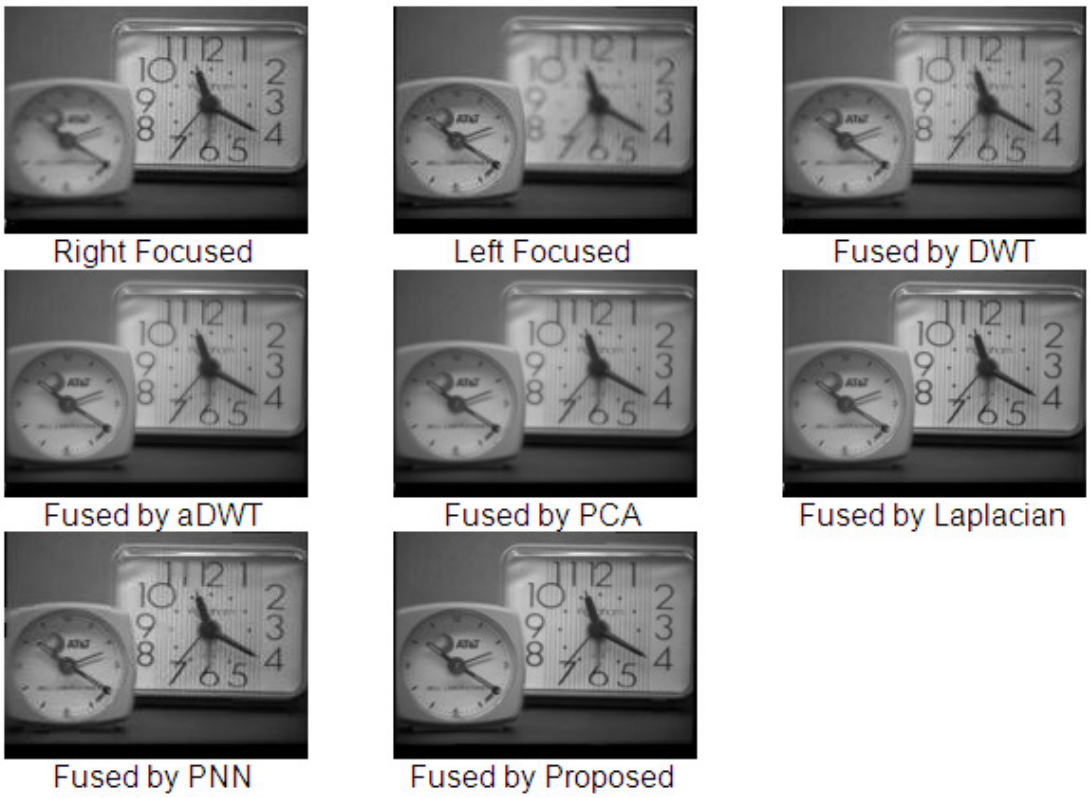

Figure 6. Clock images fused by different image fusion techniques and the proposed method.

image block. Block size is determined adaptively for each image. The trained neural network is then used to fuse any pair of multi-focus images. Experimentation results show that the proposed technique performs better than the existing techniques.

By finding the block size adaptively, the blurred and un-blurred regions within the source images are optimally identified. As a result of it, the proposed technique performs better. In the proposed technique, only one neural network is created whereas in PNN-based image fusion (Li, 2002), neural network for every pair of multi-focus images is created which is really time consuming.

Figure 7. Pepsi can images fused by different image fusion techniques and the proposed method.

Table 4. Experimentation results of different quantitative measures for clock and Pepsi can images.

Image	Method	Quality Measures			
		Entropy	SD	SF	MI
Pepsi can	DWT	7.1110	44.0932	11.6876	10.2706
	aDWT	7.1104	44.1003	11.6940	10.2801
	PCA	7.0895	44.0054	10.6239	11.6911
	Laplacian	7.1254	44.3266	12.3542	12.8991
	PNN	7.0985	45.3052	11.7414	12.9683
	Proposed	7.1343	45.5913	11.7534	13.7191
Clock	DWT	7.3710	50.7335	14.9347	6.8656
	aDWT	7.3706	50.7341	14.9256	6.8659
	PCA	7.0895	44.0054	10.6239	11.6911
	Laplacian	7.1254	44.3266	12.3542	12.8991
	PNN	7.3829	50.9907	19.5423	13.1160
	Proposed	7.3982	51.9499	22.2377	13.9956

REFERENCES

Blum RS, Zheng L (2006). Multi-sensor image fusion and its applications.CRC Press, Taylor and Francis Group, Boca Raton.

Cvejic N, Loza A, Bull D, Cangarajah N (2005). A similarity metric for assessment of image fusion algorithms. Int. J. Signal Proc., 2(3): 178-182.

De I, Chanda B (2006). A simple and efficient algorithm for multifocus image fusion using morphological wavelets. Signal Processing, pp. 924-936.

Goldberg DE (1989). Genetic Algorithm in Search Optimization and Machine Learning. Addison Wesley Publishing Company, Redwood City, Cam., 1-23.
Heijmans HJ, Goutsias J (2000). Multiresolution signal decomposition schemes Part 2: morphological wavelets. IEEE Trans. Image Process., 9: 1897-1913.
Li H, Manjunath S, Mitra SK (1995). Multi-sensor image fusion using the wavelet transform. Graphical Models and Image Processing, 57(3): 235-245.
Li S, Kwok JT, Wang Y (2002). Multifocus image fusion using artificial neural networks. Pattern Recogn. Letters, 23(8): 985-997.
Naidu VPS, Girija G, Raol JR (2003). Evaluation of data association and fusion algorithms for tracking in the presence of measurement loss. In: AIAA Conference on Navigation, Guidance and Control, Austin, USA, pp 11-14.
Naidu VPS, Raol JR (2008). Pixel-level Image Fusion using Wavelets and Principal Component Analysis. Defence. Sci. J., 58(3): 338-352.
Pajares G, de la Cruz JM (2004). A wavelet-based Image Fusion Tutorial. Pattern Recognition, 37(9): 1855-1872.
Toet A (1989). Image fusion by a ratio of low pass pyramid. Pattern Recognition Letters, 9(4): 245-253.
Wang Z, Bovik AC (2002). A universal image quality index. IEEE Signal Proc. Lett,, 9(9): 81-84.
Zhang X, Han J, Liu P (2005). Restoration and fusion optimization scheme of multifocus image using genetic search strategies.Optica Applicata, 35: 4.
Zheng Y, Essock EA, Hansen BC (2004). An Advanced Image Fusion Algorithm Based on Wavelet Transform – Incorporation with PCA and morphological Processing. In: Proceedings of the SPIE, 5298: 177-187.