ISSN 2047-3338

An all Approach for Multi-Focus Image Fusion Using Neural Network

Madhavi R.[1] and Prof. K. Ashok Babu[2]

***Abstract*-- 'Image Fusion' is the Information from multiple images which are combined to generate the new image, the generated image is more suitable for humans and machines for further image-processing tasks like image segmentation, edge detection, stereo matching, enhancement, extraction and recognition. Novel feature-level multi-focus image fusion technique is proposed in this paper, which fuses multi-focus images using classification. In this technique, Multi-focus images of ten pairs are divided into blocks and the most favorable block size for each image was found in an adaptive manner. The resultant block feature vectors are fed to feed towards neural network. The neural networks are trained in such a way that, the trained neural networks are used to fuse any pair of multi-focus images. The proposed and implemented technique used in this paper is more efficient and useful; to highlight the efficiency we have performed broad experimentation on this technique. The comparisons of the different approaches with previous methods along with the current method by calculating different parameters are also provided.**

***Index Terms*— Image Fusion, Features, Neural Networks and Block Size**

I. Introduction

IMAGE fusion is the process of combining relevant information in two or more image of scene into a single highly informative image. The image fusion method tries to solve the problem of combining information from several images taken from the same object to get a new fused image.

Image fusion applications has its own importance because these applications are used in remote sensing, medical sciences, forensic, defense and machine vision departments. Several situations in image processing require high spatial and high spectral resolution in a single image [11]. Most of the available equipment is not capable of providing such data convincingly. The image fusion techniques allow the integration of different information sources. Image fusion can be processed and performed for multi focus and multi sensor images of the same picture. In multi-focus image fusion, we considered situations where two or more images that depict the same scene will not be in-focus everywhere (if one object is in-focus, another one will be out of focus). This occurs because there are sensors which cannot generate images of all objects at various distances with equal sharpness. Examples of multi-focus images could be found in optical microscopy. Limited depth of field is a common problem with conventional light microscopy. Since the specimen is sectioned by moving the object along the optical axis, portions of the objects surface outside the optical plane will be defocused in the acquired. The advantages of multi-focus images can be fully exploited by merging the sharply focused regions into one image that will be in-focus everywhere. Multi-sensor images of the same scene are captured by different sensors where as multi-focus images are captured by the same sensor [1].

The perfect image with all the objects are in focus can be achieved with two processes of image fusion in spatial domain process pixel values [12] are directly incorporated whereas in transformed domain the images are transformed into multiple levels of resolutions [3].

There are a number of different levels in which image fusion is generally performed. Simple techniques in which the fusion operation is performed directly on the source images (e.g. weighted average method). In pixel level simple mathematical operations like max or mean (average) are applied on the pixel values of the source images to generate fused image. However, this technique usually smooth the sharp edges are leave the blurred effects in the fused image [4], [9]. 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 [8]. Using some fusion rule, the regions 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.

The fusion methods such as averaging, Brovey method, principal component analysis (PCA) [5] and IHS based methods fall under spatial domain approaches, another important spatial domain fusion method is the high pass filtering based technique. Here the high frequency details are injected into unsampled version of MS images. The disadvantage of spatial domain approach is that they produce spatial distortion in the fused image. The multi resolution analysis has become a very useful tool for analyzing remote sensing images [6]. The discrete wavelet transform has become a very useful tool for fusion [1]. Some other fusion

Madhavi R.; Author is with the Sri Indu College of Engg., Student, HYD, AP, India, (email id:madhavi.ramidi7@gmail.com)

Prof. K. Ashok Babu; Author is with the Sri Indu college of Engg,, AP, India. He is now with Designation HOD for Department of Electronics, in Sri Indu College of Engg., JNTU University, HYD, India

methods are also there, such as Lapacian pyramid based [2], curvelet transform based etc. These methods show a better performance in spatial and spectral quality of the fused image compared to other spatial methods of fusion.

II. PROPOSED METHOD

We have proposed a new method for multi-focus image fusion in this paper; the same is discussed in below sections. The performance of the proposed method was calculated using different features and the same was described in section III. Experimentation details are provided in Section IV. Section V concludes the paper.

In this proposed method, to train our neural network [13], [14], we have used ten different images. Left Focused and Right Focused versions of each image have to be created. Once these versions were created we have divided the each image into number of blocks. The block size is important here to discriminate the blurred and un-blurred sections from each other. To accomplish this task, we ran the algorithm for finding adaptive block size using Genetic Algorithm for every image in the image set. This Genetic Algorithm was used to search procedures based on mechanics of natural selection and natural genetics. Xinman Zhang et al. [7], [8] used a population of chromosomes where every chromosome represents the width and length of the block. Once the images are divided into blocks, the feature values of each block of all the images are calculated and a features file will be created.

To train the neural network an adequate number of feature vectors were used, then the trained neural network was used to fuse any set of multi-focus images. In the below section we have discussed the Image Dataset method.

A. *Image Dataset Creation*

In the proposed method, we first created an image-set of 10 grayscale images. This image-set was shown in the Fig. 1. These images are taken from different image processing websites. The images are of different scenes and backgrounds. For every image in the set, we created its two versions of the same size. In the first version, some of the regions are randomly selected in the left half of the image and are blurred. A similar process was performed in the right half of the image in the second version. The blurred versions are generated by motion blurring of Motion 10*30, 15*40, 20*50 for the image set of 10 grayscale images, we created 20 versions of blurred images.

B. *Feature Selection*

The important task in feature-level image fusion was selection of dissimilar features. Some of the objects are clear (in focus) and some objects are blurred (out of focus) in multi-focus images. Clearness of the image was reduced by blurred objects. In this proposed method we have used 5 different features to exemplify the information level contained in a specific portion of the image. This features set consists of Variance, Energy of Gradient, Contrast Visibility, spatial Frequency and Canny Edge information. From the Fig. 2, we observe how the blurriness has changed by

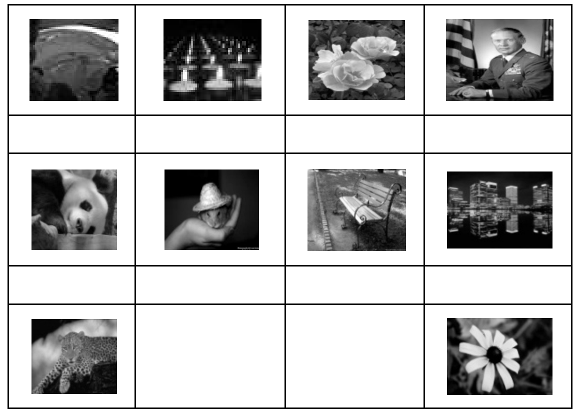

Fig. 1: Trained neural network image set

Fig. 2: Original Image along with Motion Blurred Images with different Motion dimensions

increasing Motion and affect the clearness of the image.

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 Eq. (1).

$$VI = \frac{1}{m*n} \sum_{(i,j)\epsilon B_k} \frac{|(I(i,j) - \mu_k)|}{\mu_k} \qquad (1)$$

Here μ_k and $m * n$ are the mean and size of the block B_k respectively.

2. *Spatial Frequency:* Activity level of an image was measured by spatial frequency. It was used to calculate the frequency changes along rows and columns of the image. Spatial frequency was measured using equation (2).

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

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

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

Here I is the image and m ∗ n was the image size. A large value of special frequency describes the large information level in the image and therefore it measures the clearness of the image.

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

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

Here μ is the mean value of the block image and m ∗ n is the image size. A high value of variance shows the greater extent of focus in the image block.

4. *Energy of Gradient (EOG):* It was also used to measure the amount of focus in an image. It was calculated using equation(4)

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

Here m and n represents the dimensions of the image block a high value of EOG 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 one if the current pixel belongs to some edge in the image otherwise it returns zero. The edge feature was just the number of edge pixels contained within the image block.

C. Proposed Algorithm

1. Find the optimal block size for each set of LF_i and RF_i using [3]. LF_i has the left focused and RF_i has the right focused versions of the *ith* image in the dataset discussed in II.a. Here i=1,2,3,…10.
2. Divide the versions LF_i and RF_i of every image in the dataset into *K* number of blocks of size M*N.
3. Create the features file for all LF_{ij} and RF_{ij} according to the features discussed in II.b. Here *j*=1,2,3,…K for 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].
4. Assign the class value to every block *j* of *ith* image. If block *j* was 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
5. 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.
6. By using the trained neural network, identify the clearness of all the blocks of any pair of multi-focus images to be fused.
7. Fuse the given pair of multi-focus images block by block according to the classification results of neural network such that

$$\text{Output of NN for block } J \begin{cases} \text{if } > 0, \text{ Select } j \text{ from left}-\text{focused image} \\ \text{if } < 0, \text{ Select } j \text{ from right}-\text{focused image} \end{cases}$$

III. Quantitative Measures

There are different quantitative measures which were used to evaluate the performance of the fusion techniques. We used six measures Entropy, Correlation coefficient, Standard Deviation, PSNR (Peak Signal to Noise ratio), RMSE (Root Mean Square Error), and MAE (Mean Absolute Error) when the reference image was available. For blind image fusion (reference image was not available) we have used Entropy, Standard Deviation, Correlation coefficient and Mean Absolute Error (MAE).

IV. Experiments and Results

Image fusion was performed in two different situations in first situation the reference image was available and in second situation, the reference image was not available (blind image fusion). We exercised the feet 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 is kept as 0.01. The results of proposed technique are compared with different existing methods including DWT, ADWT and PCA based image fusion techniques. The experimentation results are obtained when the reference image was available and when it was not available (blind image fusion). In order to evaluate the performance of the proposed technique the result for 4 different pairs of multi-focus images are obtained including Rice, Pepsi, Beach and Clock.

A. Variation between the proposed technique and PNN based technique

Shutao Li et al. 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 image to generate the fused image. In their technique the source were divided into the blocks of fixed

TABLE 1: PERFORMANCE METRICS USED FOR IMAGE FUSION

Parameter	Equation	Description
CORRELATION COEFFICIENT	$r_K = \frac{\sum_{t=1}^{N}(x_t - \overline{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{N}(x_t - \overline{x})^2}$ Where x_t is a data value at time step t,k is the lag.	The correlation coefficient matrix represents the normalized measure of the strength of linear relationship between variables.
STANDARD DEVIATION	SD=$\sqrt{\sum_{i=0}^{L}(i-i')^2 h_F(i)}$ Where $h_F$the normalized histogram of fused image and L is is the number of grey levels.	SD will measure the contrast in the fused image. Standard deviation will be high for a well contrast image.
MI	$MI = \sum_{i=1}^{m}\sum_{j=1}^{n} h_{R,F}(i,j)\log_2\left[\frac{h_{R,F}(i,j)}{h_R(i,j)h_F(i,j)}\right]$ Where $h_{R,F}$, h_R, h_F are the joint, reference and fused images histograms respectively.	Determines the information of fused image retrieved from the input source images. A higher value shows good quality results.
PSNR	$PSNR = 20\log_{10}\left[\frac{L^2}{\frac{1}{m*n}\sum_{i=1}^{m}\sum_{j=1}^{n}[R(i,j) - F(i,j)]^2}\right]$ Where L denotes the number of grey levels in the image.	Determines the degree of resemblance between reference and fused image. A higher value shows good quality results.
RMSE	$RMSE = \sqrt{\frac{1}{m*n}\sum_{i=1}^{m}\sum_{j=1}^{n}[R(i,j) - F(i,j)]^2}$ Here R,F are the reference and fused images respectively. $m*n$ Is the image size.	Calculates the deviation between pixel values of reference image and fused image. A smaller value shows good quality results.
ENTROPY	$H = -\sum_{i=0}^{L-1} h_F(i)\log_2 h_F(i)$ Where $h_F$the normalized histogram of fused image and L is is the number of grey levels.	Quantifies the quantity of information contained in the fused image. A bigger value shows good quality results.
MAE	$MAE=\frac{1}{N}\sum_{i=1}^{n}\lvert f_i - y_i\rvert$ Where $f_i$the predicted is fused image and y_i is the true value fused image.	It calculates the error between true value and predicted value.

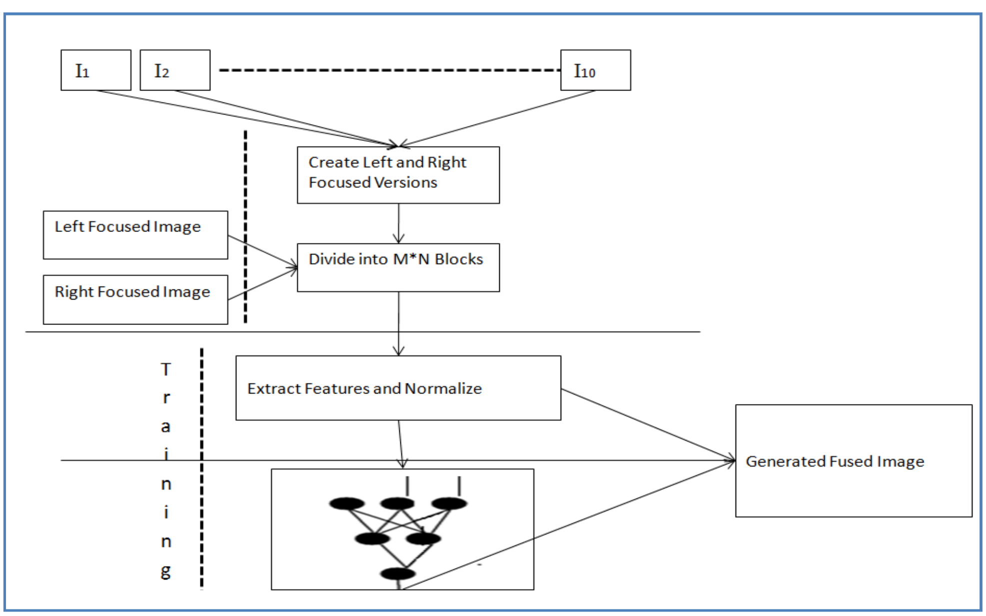

Fig. 3: Block diagram of the proposed method

size of 32*32. The block size has an important factor to achieve good fusion results. Fixed sized blocks may separate the blurred and unblurred 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 image to image because different images have different blurred region. In the Proposed method for every pair of multi-focus images a optimal block size is found using the technique given in [3]. We have used five different features to calculate the clearness of a block more accurately as compared to three features used by Shutao Li et.al.

A major difference between the proposed method and existing PNN Based image fusion technique was the training of the neural network Shutao Li et.al. In their technique create and train a new neural network for every pair of multi-focus images which was really 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 was obtained then it can be use to fuse any pair of multi-focus images.

B. Quantitative Assessments and the Visual Comparison

When the reference image is available: For Rice and Pepsi images the reference images are available. A visual comparison has shown in figures 4 for Rice and Pepsi images.

When the Reference Image is not available (Blind Image fusion): We have used Beach and Clock Images in this category. The clock and Beach image sizes are 256*256. When the reference image was not available then the performance of the fusion process was evaluated on the basis of different set of quantitative measures are described in the Table 3. Fig 5 provides a visual comparison of the proposed technique with the existing techniques.

V. CONCLUSION

In this proposed technique, a feature level block-based multi-focus image fusion has proposed in this paper. In this method we have trained the feed forward neural network with the block features of 10 pairs of multi-focus images. A feature set including spatial frequency, contrast visibility, edges, variance and energy of gradient is used to define clarity of the image block. Here adaptive block size was determined for each image adaptively. The trained neural network was 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 was created whereas in PNN-based image fusion, neural network for every pair of multi-focus images was created which was really time consuming.

REFERENCES

[1] H.Li, S. Manjunath and S.K. Mitra, “Multi-sensor image fusion using the wavelet transform” in Graphical Models and Image Processing, vol. 57, 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] Ishita De and Bhabatosh Chanda, “A simple and efficient algorithm for multifocus image fusion using morphological wavelets” in Signal Processing. pp. 924-936, 2006.

[4] 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.

[5] 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.

[6] H.J. Heijmans and J. Goutsias, “Multiresolution signal decomposition schemes”, Part 2: “morphological wavelets”, in IEEE Trans. Image Processing 9, pp. 1897-1913, November 2000.

[7] Xinman Zhang, Jiuqlang Han and Peifei Liu, “ Restoration and fusion optimization scheme of multifocus image using genetic search strategies”, Optica Applicata, Vol. XXXV, No. 4, 2005.

[8] 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 Proceedings of the SPIE, vol 5298, pp. 177-187, 2004.

[9] MIAO Qiguang, WANG Baoshu,Rahman Ziaur Schowengerdt Robert A.” Multi-focus image fusion using ratio of blurred and original image intensities” Proceedings of SPIE, the International Society for Optical Engineering . Visual information processing XIV : (29-30 March 2005, Orlando,Florida, USA) .

[10] Lahouari Ghouti, Ahmed Bouridane and Mohammad K. Ibrahim,” Improved Image Fusion Using Balanced Multiwavelets.

[11] Wei-Wei Wang; Peng-Lang Shui; Guo-Xiang Song “Multifocus image fusion in wavelet domain” Machine Learning and Cybernetics, 2003 International Conference on Volume 5, Issue , 2-5 Nov. 2003 Page(s): 2887 – 2890 Vol.5.

[12] Gemma Piella” A general framework for multiresolution image fusion: from pixels to regions” Information Fusion Volume 4, Issue 4, Decembe 2003, Pages 259-280.

[13] Shutao Li, James T. Kwok, Yaonan Wang “Multifocus image fusion using artificial neural networks” Pattern Recognition Letters Volume 23, Issue 8 (June 2002) Pages: 985 - 997 Year of Publication: 2002 ISSN:0167- 8655 .

[14] Wei Huang, Zhongliang Jing “Multi-focus image fusion using pulse coupled neural network” Pattern Recognition LettersVolume 28 , Issue 9 (July 2007) Pages: 1123-1132 Year of Publication: 2007.

Pagidimarry Madhavi, graduated from Nirmal Engineering College, Mancherial in Electronics & Communications Stream. Now pursuing Masters in Digital Electronics and Communication Systems (DECS) from Sri Indu College of Engineering & Technology.

Dr. K. Ashok Babu, Professor & Head of the Department of (ECE) and for his constant co-operation, support and for providing necessary facilities throughout the M. Tech program. He has 15 Years of Experience, at B.Tech and M. Tech Level and he was working as a Professor in Sri Indu College of Engg. and Technology.

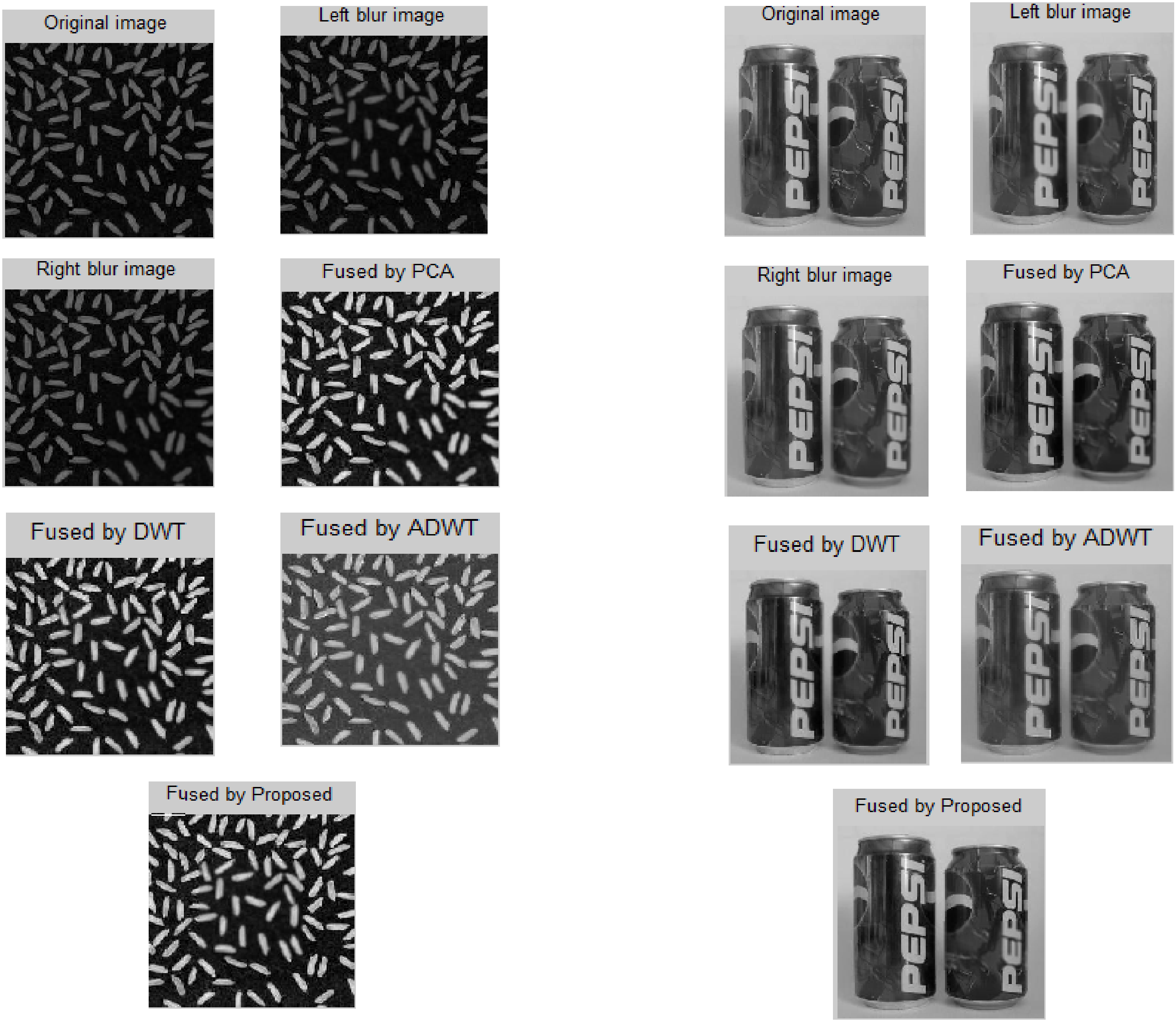

Fig. 4: Rice and Pepsi Images fused by different image fusion techniques and the proposed method.

TABLE 2: RESULTS OF QUANTITATIVE MEASURES FOR RICE AND PEPSI IMAGES.

Parameter	**Rice Image**				**Pepsi Image**			
	DWT	ADWT	PCA	Proposed	DWT	ADWT	PCA	Proposed
Correlation coefficient	0.9863	0.8522	0.9801	0.9826	0.9952	0.9457	0.9892	0.9914
Entropy	0	0.0088	0.1274	0.1442	0.9979	0.9944	0	0
Standard Deviation	0.1318	0.2554	8.4734e+003	8.4565e+003	0.1654	0.3213	1.0614e+004	1.0551e+004
PSNR	14.6300	14.6300	31.6006	31.9836	5.8236	5.8236	29.8182	30.7096
RMSE	47.3196	47.3197	6.7067	6.4174	130.4245	130.4245	8.2344	7.4312
MAE	32.3773	32.3773	0.0456	0.0288	117.4573	117.4574	0.0754	0.0329

TABLE 3: RESULTS OF QUANTITATIVE MEASURES FOR BEACH AND CLOCK IMAGES.

Parameter	**Beach Image**				**Clock Image**			
	DWT	ADWT	PCA	Proposed	DWT	ADWT	PCA	Proposed
Correlation coefficient	0.9904	0.9394	0.9812	0.9914	0.9937	0.9566	0.9891	0.9893
Entropy	0.9493	0.9466	0.0318	0.0606	0.6952	0.7048	0	0
Standard Deviation	0.2720	0.5338	1.7438e+004	1.7594e+004	0.1717	0.3394	1.0979e+004	1.0957e+004
MAE	94.6210	94.6210	0.0867	0.0106	186.2614	194.7580	0.1199	0.1481

Fig. 5: Beach and Clock Images fused by different image fusion techniques and the proposed method