

Spatial Stimuli Gradient Based Multifocus Image Fusion Using Multiple Sized Kernels

Muhammad MUZAMMIL*, Imdad ALI, Umer JAVED, Muhammad AMIR, Ihsan UIHAQ

Abstract: Multi-focus image fusion technique extracts the focused areas from all the source images and combines them into a new image which contains all focused objects. This paper proposes a spatial domain fusion scheme for multi-focus images by using multiple size kernels. Firstly, source images are pre-processed with a contrast enhancement step and then the soft and hard decision maps are generated by employing a sliding window technique using multiple sized kernels on the gradient images. Hard decision map selects the accurate focus information from the source images, whereas, the soft decision map selects the basic focus information and contains minimum falsely detected focused/unfocused regions. These decision maps are further processed to compute the final focus map. Gradient images are constructed through state-of-the-art edge detection technique, spatial stimuli gradient sketch model, which computes the local stimuli from perceived brightness and hence enhances the essential structural and edge information. Detailed experiment results demonstrate that the proposed multi-focus image fusion algorithm performs better than the other well known state-of-the-art multifocus image fusion methods, in terms of subjective visual perception and objective quality evaluation metrics.

Keywords: focus map; image fusion; multi-focus; multi kernel; spatial domain fusion

1 INTRODUCTION

Usually, imaging cameras have the limitation of a finite field of depth which causes a partially focused scene acquired from the optical lens [1]. The objects located within the field of depth of a camera have sharp details, whereas the rest of the objects are blurred [2]. Partially focused images often provide limited performance in various applications including surveillance, remote sensing, medical imaging and object recognition [1]. Therefore, multiple images are acquired and the complementary information of these images is combined into one single image by using image fusion techniques.

Multifocus image fusion algorithms can be broadly classified into two categories: transform domain algorithms [3] and spatial domain algorithms [4]. A sub-category of transform domain fusion algorithms is the multiscale transform. Multiscale transformation is applied on the source images to obtain multi-resolutions and decomposed coefficients of low and high frequencies or orientations. Based upon the fusion rule, different sub-band coefficients are considered, and the final fused image is obtained by taking inverse multiscale transform. Early proposed multiscale transform fusion methods include: pyramid decomposition [5], wavelet transform [6], complex wavelet transform [7] and contourlet transform [8]. Some of the latest multiscale transform fusion methods are proposed in [9-12]. Since these kinds of algorithms apply a global fusion technique, hence, for misregistered images they produce poor results [13]. Moreover, These methods generally produce a low contrast fused image [14].

Spatial domain algorithms usually find the saliency map/focus regions in the source images [14] for fusion. Block based fusion methods have been used widely because the focus measure cannot be efficiently represented by a single pixel [15]. The optimum size of block is another challenge in these types of methods because the larger block size may result in degraded weight calculation, especially for a block having both, focused and unfocused areas.

Usually spatial domain fusion algorithms are immune to shift variance and due to simplicity, they are the best choice for incorporation in real-time devices. Saliency detection for focus measure is the basic requirement for

spatial domain algorithms. Normally, gradient information is used to detect the focused areas in the image. Scale Invariant Feature Transform (SIFT), is a well-known algorithm to find the feature vectors that contain gradient information on key points in an image [16]. However, dense SIFT (DSIFT) finds 128D feature vector for each pixel in the image, in contrast to SIFT which finds 128D feature vector for key points only [17]. Baseline for DSIFT algorithms is the binning of magnitude and directions of the gradients. Therefore, Liu et al. used DSIFT for determination of activity map for multi-focus image [18]. The DSIFT can be used for the registration of misaligned source images; however, most of the available datasets in case of multi-focus image fusion are pre-registered. Moreover, activity map determination using DSIFT is computationally expensive as compared to some latest gradient based implementations such as; Spatial Stimuli Gradient Sketch Model (SSGSM) proposed by Mathew and James [19].

In this paper, we used a multiple sized kernel technique to determine the focus map. A kernel size of n and $n \times n$ is used simultaneously to measure the focus activity in each of the source images. The kernel size of n and $n \times n$ determines the hard and soft decision maps, respectively. The hard decision map has the tendency to differentiate between the boundary of focused and unfocused regions, whereas the soft decision map strikes out the outliers falsely detected by the soft decision maps. The focus map is finally determined by combining the hard and soft decision maps. Histogram equalization is applied as a preprocessing step to enhance the edge information in the source images.

The remainder of the paper is organized as follows: section 2 describes the detailed fusion scheme of the proposed algorithm. Experimental setup and results are discussed in section 3. Finally, conclusions are drawn in section 4.

2 PROPOSED FUSION SCHEME

The flow diagram of our proposed fusion scheme is given in Fig. 1. As most of the datasets available for multi-focus image fusion are preregistered, we assume that both the source images are preregistered in this work. Smooth regions of the source image belong to unfocused area and

the region containing sharp edges from a part of the focused area. Therefore, for pre-processing, contrast of both the source images is enhanced by using the Non-parametric Modified Histogram Equalization (NMHE) presented in [20]. This helps boost the information regarding the edges available in the source images. The fusion scheme proposed in this paper is mainly divided into three steps. First, local stimuli map of both the pre-processed source images is calculated by using SSGSM. Local stimuli maps are used to compute the activity level maps which contain the information of focused areas in both the images. Following that, the determined focused areas of both source images, along with the undetermined area are computed in the coarse decision map by using the focus information in both activity level maps. The next step is to refine the undetermined area of the coarse decision map by using the local focus measure to obtain a final decision map. As a last step, the final decision map is used to obtain the fused image.

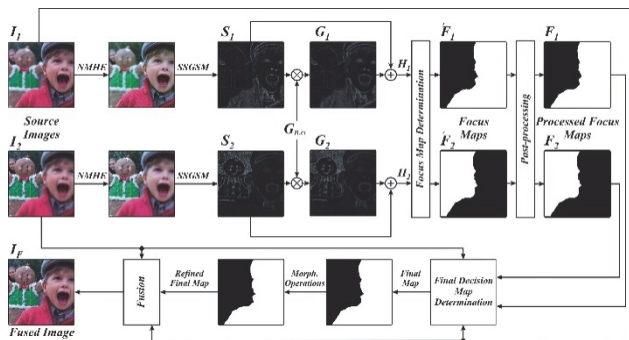


Figure 1 Flow diagram of proposed fusion algorithm

Histogram equalization is widely used to enhance the low contrast images so that the original image can be mapped as close as possible to the uniform distribution of the intensities in the histogram. NMHE is one of the histogram equalization methods which not only enhances the contrast, but also preserves the average brightness of the original image.

We used NMHE as a pre-processing step for the contrast enhancement of source images. NMHE algorithm calculates an updated histogram by removing the spikes from the original histogram whose occurrence in the histogram is very high as compared to neighbouring intensity levels. The first step of NMHE suppresses the spikes from original histogram. Afterwards the algorithm clips and normalizes the histogram of the previous step and then calculates the cumulative deviation of the transitional updated histogram from the uniform histogram. It then uses this as a weighting factor to construct a final updated histogram that is a weighted mean of the updated histogram and the uniform histogram. It then uses this as a weighting factor to construct a final modified histogram that is a weighted mean of the modified histogram and the uniform histogram.

Fig. 2 shows the change in gradients after applying NMHE on "joy", "flower", "clock" and "toy" source images, respectively. Fig. 2a-d show the source images and their gradients are shown in Fig. 2e-h, whereas, Fig. 2i-l show the contrast stretched source images and their gradients are shown in Fig. 2m-p. It can be observed that Fig. 2m-p have a higher gradient magnitude as compared

to the gradient magnitudes of the original source images shown in Fig. 2e-h.

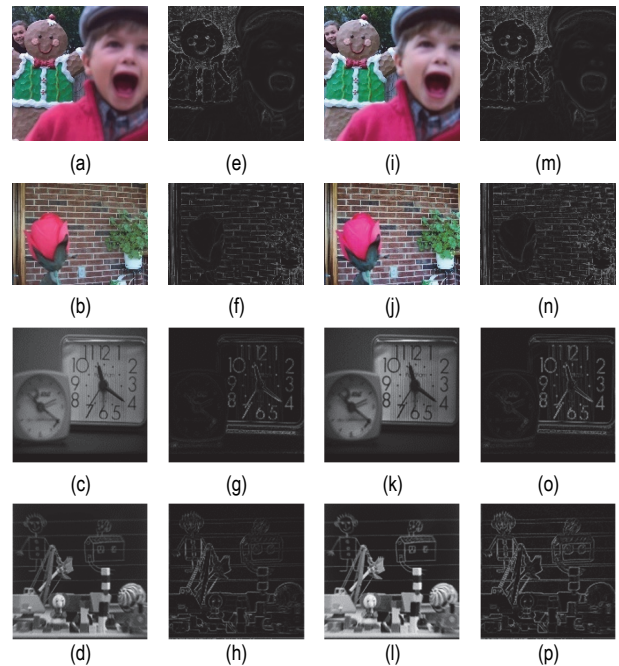


Figure 2 Effect of preprocessing technique using NMHE, a)-d) Original images; e)-h) Gradients of a)-d) images respectively; i)-l) Contrast stretched using NMHE of a)-d) images respectively and m)-p) gradient of i)-l) images respectively

Eq. (1) determines a threshold and the pixels higher than the threshold contribute in the modified histogram.

$$h_m(n) = p[n | K] \quad (1)$$

where, $p[n | K]$ is the occurrence probability of the n th intensity level given the horizontal variation of contrast C . A measure of equalization (M_{eq}) indicates the non-uniformity of histogram distribution of an image, which is calculated according to Eq. (2):

$$M_{eq} = \text{sum}(p_u - h_{mc}) \quad (2)$$

where, p_u is uniform probability density function and c is the modified clipped histogram calculated from original histogram. M_{eq} assigns the weight to absolute uniform and the modified histogram to calculate the redesigned histogram (h_{NMH}) as per the following Eq. (3):

$$h_{NMH} = (1 + M_{eq})h_{mc} + (M_{eq})p_u \quad (3)$$

From CDF of h_{NMH} , transformation curve $T(n)$ is obtained Eq. (4):

$$T(n) = \lfloor (L-1)c(n) + 0.5 \rfloor \quad (4)$$

where, $c(n)$ is the CDF of the h_{NMH} and $L = 256$ for an 8-bit image. $T(n)$ is applied to the original image to obtain the contrast improved images. Contrast enhancement results in improved edges in the source images.

The next step to the fusion scheme is to identify the focused areas. Various gradients based techniques are

proposed in the literature like DSIFT, SSGSM and local standard deviation (LSD). LSD computes the deviation in a neighbourhood of 3×3 , whereas, DSIFT finds a feature vector of 128D for each pixel of an image which contains the 8 orientation bins of gradients in a grid of 4×4 . The SSGSM is based on two well-known laws: Weber-Fechner law of perceived brightness and Sheperd similarity law pertaining to neighbourhood similarity [19]. The perceived brightness B_p of an image I is evaluated from following Eq. (5):

$$B_p = a \log_{10}(I) \quad (5)$$

where, a is the constant of proportionality.

Gradients represent the edges in an image; however, the gradient is a linear operator so it cannot suppress the noise available in the images. Moreover, gradient measures the distance from the local intensity variability point of view. One of the methods proposed by [19] is the measure of perceived similarity capable of screening out the noise edges because of the exponential transformation. Dissimilarity of B_p along x and y axis are calculated as D_x and D_y , respectively, and the function of corresponding gradients G_x and G_y are given in Eq. (6) below:

$$D_i = G_i \left(e^{-G_i^2} \right); \forall i; i \in \{x, y\} \quad (6)$$

Eq. (7) gives the magnitude of local stimuli D as follows:

$$D = \sqrt{D_x^2 + D_y^2} \quad (7)$$

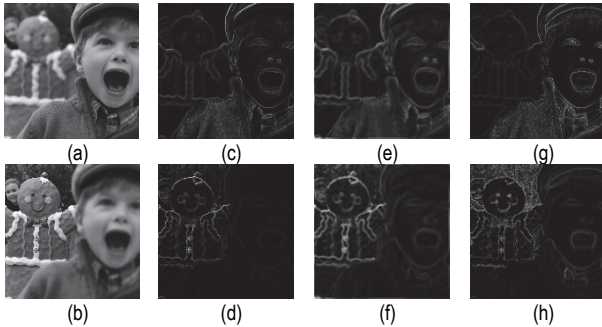


Figure 3 Comparison of gradient features, a) Near focused greyscale source image; b) Far focused greyscale source image; c-d) LSD of near focused and far focused greyscale images, respectively; e-f) DSIFT and g-h) SSGSM near focused and far focused grayscale images, respectively

Fig. 3c-d shows the results after applying LSD on both the grayscale source images shown in Fig. 3a-b and Fig. 3e-f shows the results after applying DSIFT and finally Fig. 3g-h shows the results after applying SSGSM on grayscale source images. The SSGSM has computed the gradient more accurately which can be observed by carefully observing the leaves in far focused SSGSM image and in the cloths of boy in near focused SSGSM image. Once the local stimuli are calculated for both the source images, the Gaussian averaging filters G_1 and G_2 with a mean value of 11 and σ of 5 are applied to the outputs of both SSGSM. The averaging filter smoothens the sharp edges recovered

in the previous step, to avoid strict boundary decision in the forthcoming step. The output of the Gaussian averaging filter is denoted by G_1 for the first source image and G_2 for the second source image.

Next step is the determination of basic focus maps. We achieved this by finding the definite focus regions of the first source image and the definite focus regions of the second source image.

Similarly, the second basic focus map contains the definite focus regions of the second source image and definite focus regions of the first source image. The detailed scheme is described in the following four stages:

Stage 1: We used $n \times n$ sliding window approach to make the fusion process shift-invariant. The size of n is selected as 3 for window W and $n \times 3$ for window W' . These two window sizes are chosen to determine hard and soft decision maps, respectively. Furthermore, windowing operation is also used to reduce the blocking artefacts in the coarse decision maps D_1 and D_2 . Step size for both pairs of sliding windows is set to one pixel. For a source image with the resolution of $P \times Q$, total number of sliding window patches for window W is: $h = (P - n + 1) \times (Q - n + 1)$ and these sliding windows patches are denoted by W_1^h and W_2^h . Similarly, the patches for window W' are denoted by $W_1^{h'}$ and $W_2^{h'}$. Then create two pairs of score matrices M_1 , M_2 and M_1' , M_2' . The first pair of score matrices i.e. M_1 and M_2 stores the focus information of near and far focused source image, respectively, computed through window size W . Similarly, M_1' and M_2' stores the focus information of near and far focused source image respectively, computed through the window W' . These score matrices are initialized with zero value and have the same size as the source images in order to store the pair of near and far focus maps acquired through windows W and W' . The score matrices are further used to determine the focus map against each pixel from two source images, as shown in Fig. 4.

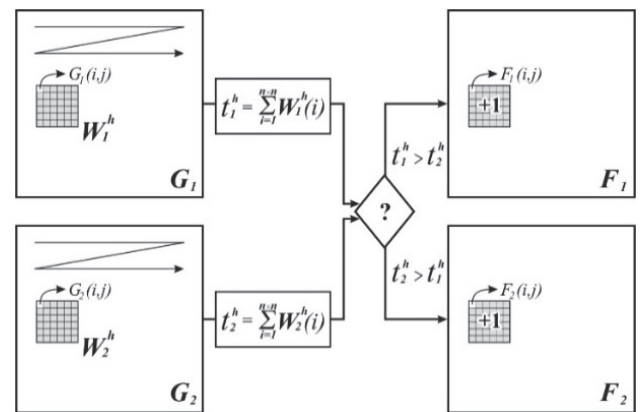


Figure 4 Flow diagram for determination of focus map

For a pair of sliding windows W_1^h and W_2^h , calculate the total focus values by summing the intensity values accompanied by the corresponding window in H_1 and H_2 . H_1 and H_2 are determined from the following Eq. (8):

$$H_i = G_i + S_i \quad (8)$$

where, $i = \{1, 2\}$, S_i is the SSGSM of source images and G_i is the result of Gaussian filter with a mean value of 11 and a standard deviation of 5.

Let the focus score be denoted by t_1^h and t_2^h . If $t_1^h > t_2^h$, then all the coefficients of the corresponding window W_1^h in M_1 are incremented by 1. Similarly, if $T_1^h > T_2^h$ then all the coefficients of the corresponding window W_2^h in M_2 are incremented by 1. The same procedure is repeated to calculate the M'_1 and M'_2 for the sliding window W' . Following this scheme, focus measure score is computed for all the corresponding pixels of both the source images. Detailed flow of the focus map determination is shown in Fig. 5.

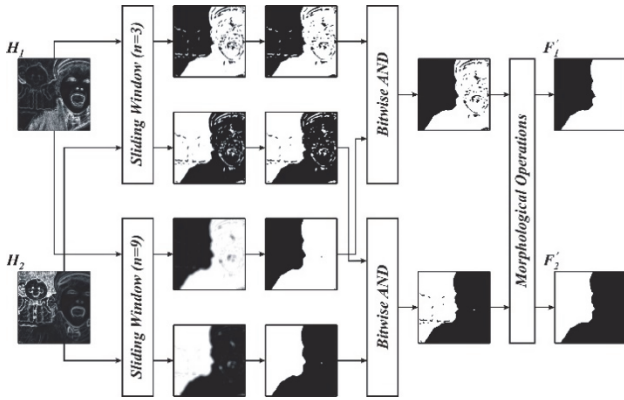


Figure 5 Flow diagram for determination of focus map

Stage 2: Each pixel in the focus maps F_1 and F_2 , calculated in the previous step is categorized as focused, defocused and undetermined pixel. If the focus measure map of a pixel (i, j) is focused in the first source image, the corresponding pixel value in the $F_2(i, j)$ must be zero. Similarly, if the focus measure map of pixel (i, j) is defocused in the first source image, the corresponding pixel value in the $F_1(i, j)$ must be zero. The reverse is the case for focused and defocused pixels of the second source image. For the remaining case, where a pixel in any focus maps $F_1(i, j)$ and $F_2(i, j)$ is non-zero, it will be declared as unclassified pixel.

Above classification rule can be summarized in Eq. (9), as follows:

$$I_n(i, j) \begin{cases} \text{focused \& defocused,} & \text{if } F_n(i, j) = 0 \\ \text{Unclassified pixel,} & \text{otherwise} \end{cases} \quad (9)$$

where, $n = \{1, 2\}$. Fig. 6 shows the classification results acquired through equation Eq. (8). Fig. 6a-b shows the near and far coarse focus maps F'_1 and F'_2 , Fig. 6c-d shows the post-processed coarse focus maps and Fig. 6e shows the absolute difference of Fig. 6c-d. The black area along the boy shows the unclassified pixels, and they need to be further processed in the next step.

Stage 3: Next step is to identify the unclassified pixels. From above classification rule, it is clear that if a pixel

$I_1(i, j)$ is focused in source image I_1 , then the same pixel must be unfocused in the source image I_2 .

The coarse decision maps obtained through above classification rule have strong tendency to segregate the focused and unfocused areas in both the source images, which can be observed in Fig. 6. Fig. 6a-b shows the coarse decision maps of source images obtained from the Fig. 3g-h respectively. Afterwards, morphological operations are applied as a post processing step on the coarse decision maps to fill the small holes existing within the focused areas and remove the small focused objects surrounded by unfocused area. Fig. 6c-d shows the images obtained after morphological operations. Fig. 6e shows the absolute difference of binary coarse decision maps shown in Fig. 6c and Fig. 6d. Black boundary around the body of boy in Fig. 6e belongs to the unclassified pixels.

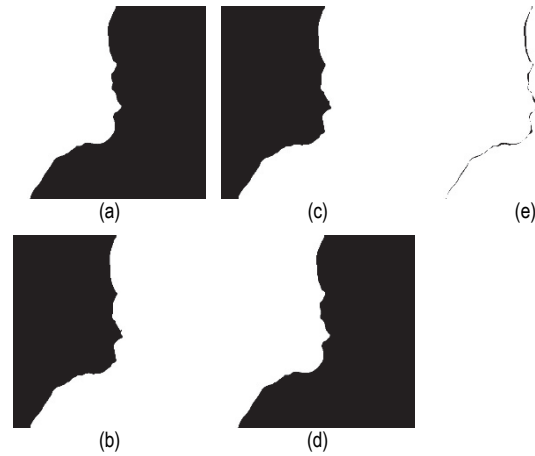


Figure 6 Coarse classification results of "joy" source images, a)-b) coarse maps; c)-d) post processing results and e) absolute difference of both decision maps

Stage 4: Final step before the fusion is to further refine the unclassified pixels in the last step. These unclassified pixels normally belong to the boundary between the focused and unfocused regions. A local descriptor, spatial frequency is used to measure the local focus of the unclassified pixel [21]. Spatial frequency can be calculated as per the following Eq. (10):

$$SF = \sqrt{RF^2 + CF^2} \quad (10)$$

where, RF is the row frequency and CF is the column frequency, which can be calculated from Eq. (11) and Eq. (12) respectively.

$$RF = \sqrt{1/(N_1 \times N_2) \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} [I(n_1, n_2) - I(n_1, n_2 - 1)]^2} \quad (11)$$

$$CF = \sqrt{1/(N_1 \times N_2) \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} [I(n_1, n_2) - I(n_1 - 1, n_2)]^2} \quad (12)$$

where, N_1 and N_2 is the total number of pixels in rows and columns of the source image I . The higher the value of spatial frequency, the higher the focus measure of the corresponding pixel.

3 EXPERIMENTAL SETUP

For detailed experiments, we have utilized 18 pairs of popular source images which are shown in Fig. 7. From these pairs of multi-focus images, 10 images are greyscale and 8 pairs are colour images. These test images are also used in [34] and a few of them are taken from the

Lytromulti-focus dataset [22]. Near focused images are shown on the left side, whereas the right side images are the far focused.

To prove the effectiveness of our proposed multi-focus image fusion algorithm, we have compared our results with the latest state-of-the-art algorithms.



Figure 7 Multi-focus test images used during the experiments

These include DSIFT [18], Discrete Wavelet Transform and Adaptive Block (DWT) [23], Guided Filter Fusion (GFF) [24], Discrete Cosine Transform (DCT) based fusion [25], Quadtree-based Multi-focus Image Fusion (QMIF) [13], Energy of Laplacian based DCT (DCT-EoL) [3] and boosted random walks-based algorithm with Two-Scale Focus maps (TSF) [26]. The proposed algorithm when compared to the DSIFT implementation has the following main improvements: a) A multiple sized kernel scheme has been introduced to determine and hard and soft decision maps. Hard decision map differentiates between the boundary of near and far

focus areas, whereas, the soft decision map reduces the falsely detected focused/unfocused regions from the final focus map. b) The dense SIFT used to calculate the activity level map does not identify the coarse decision map properly owing to the fact that dense SIFT detects many redundant features, therefore, we have used SSGSM to determine gradient features. c) Before calculating the decision maps, contrast and edge enhancement scheme has been incorporated to support the onward image fusion process.

The proposed algorithm and the algorithms used for comparisons are executed on Intel® CORE™ i5 3210M

Processor, 8 GB RAM, 64-bit Windows 10 laptop. All the algorithms were implemented on MATLAB R2014a (8.3.0.532) 64-bit version.

4 RESULTS AND DISCUSSION

4.1 Qualitative Comparison

Multi-focus source images contain both focused and unfocused areas. Hence, a fusion algorithm with the tendency to accurately detect and merge the focused areas from the source images will yield a better visual perception. For qualitative comparison, we evaluate the visual quality of the fusion results shown in Fig. 8. These are the results of different fusion techniques for a pair of pre-registered "Joy" source images shown in Fig. 7p. Three different areas from the boundary of near and far focused source images are selected and their magnified version is displayed above each fusion result for the purpose of detailed visual comparison of all the fusion techniques compared during the experiment.



Figure 8 Fusion result of "Joy" multi-focus test image, a) QMIF [13]; b) DCT [25]; c) DWT [23]; d) GFF [24]; e) DSIFT [18]; f) DCT-EoL [3]; g) TSF [26] and h) Proposed

It can be seen that GFF and DSIFT have better fusion results as compared to the rest of the fusion methods, but our proposed method has picked-up the focus regions from both the source images accurately. Adequate blur can be seen in all three areas of QMIF, DCT and DWT based fusion schemes. For the DCT-EoL fused image, numerous blocking artefacts are present. However, in case of GFF and TSF, the magnified area on the right hand side has very clear focus, but the magnified areas show the blur at the boundary of focused and unfocused areas. Similarly, for the DSIFT fusion results, the magnified area showing cap

view, but the magnified area shown on the right side contains some of the portion of unfocused part. In our proposed fusion scheme, it can be observed that all the three magnified areas contain sharp details of both near and far focused source images.

Multi-focus source images contain both focused and unfocused areas. Hence, a better fusion algorithm will have the capability to acquire the focused areas from both the source image more accurately. To further evaluate the perceptive quality of the proposed fusion algorithm, the difference between the fused and far focused "Joy" source image is shown in Fig. 9.



Figure 9 Difference of fused image and far focused "Joy" source images, a) QMIF [13]; b) DCT [25]; c) DWT [23]; d) GFF [24]; e) DSIFT [18]; f) DCT-EoL [3]; g) TSF [26] and h) Proposed

It can be observed that the Quadtree, DWT, DCT, DCT-EoL and DSIFT based fusion schemes have acquired few areas of near focused source image as well, which belongs to the unfocused area of far focused source image. TSF also acquired some of the unfocused portion into the fused image from the far focused source image. It is clear from the magnified areas of the proposed fusion results that the proposed fusion scheme has outperformed the rest of techniques in term of preceptive quality.

Fig. 10 shows the fusion results of "Clock" source images. It can be observed that the proposed fusion algorithm has clearly separated the boundaries of near and far focused images, whereas the rest of the fusion scheme has acquired some of the out-of-focus regions near the boundaries.

To assess the contribution of each of the source image in the fused image, we have used Gradient Magnitude Similarity Deviation (GMSD) [27]. Fig. 11c-d show which portion of the near and far focused source images is present in the fusion result of "clock" test images.

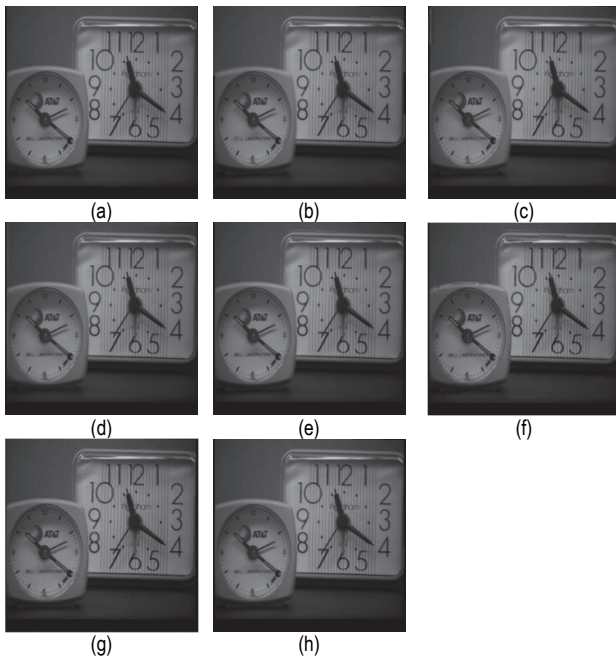


Figure 10 Fusion results of "Clock" source images, a) QMIF [13]; b) DCT [25]; c) DWT [23]; d) GFF [24]; e) DSIFT [18] and f) DCT-Eol [3]; g) TSF [26] and h) Proposed

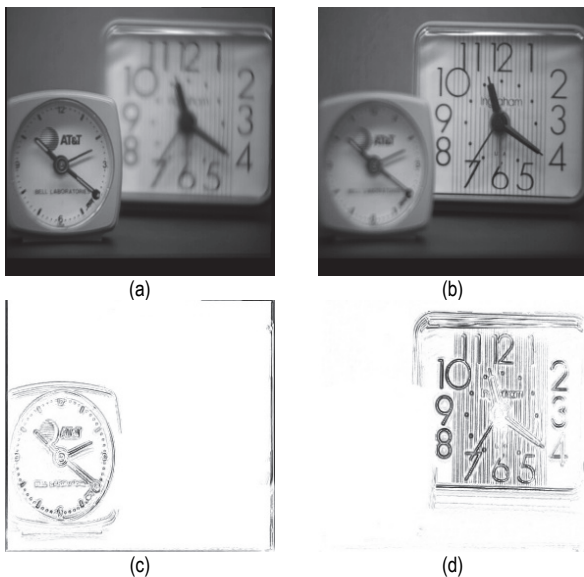


Figure 11 Contribution of near and far focused images in fused image, computed through GMSD, a)-b) near and far focused source images; c)-d) contribution of relevant source images in fused image calculated through GMSD

4.2 Quantitative Comparison

Evaluation of algorithms on the basis of visual quality is insufficient and requires a quantitative comparison. Various fusion quality evaluation metrics have been proposed, but none of them can completely evaluate the fusion quality independently [28]. Some well-known criteria for objective evaluation of fusion quality are explained as follows:

4.2.1 Spatial Structural Similarity (SSS) $Q^{AB/F}$

SSS is an edge based fusion quality evaluation metric proposed by Xydeas and Petrovic [42, 43]. This metric finds the amount of edge information transferred into the fused image from all the source images. $Q^{AB/F}$ for two

source images can be calculated from the following Eq. (13):

$$Q^{AB/F} = \frac{\sum_{x=1}^X \sum_{y=1}^Y (Q^{AF}(x,y)W^A(x,y) + Q^{BF}(x,y)W^B(x,y))}{\sum_{x=1}^X \sum_{y=1}^Y (W^A(x,y) + W^B(x,y))} \quad (13)$$

where, $Q^{FB}(x,y)$ represents the information transferred from source image A into the fused image F for the pixel location (x,y) and $W^B(x,y)$ is the weight for a particular pixel location (x,y) . A pixel with higher gradient value influences more the $Q^{AB/F}$ than the lower gradient value.

Thus $W^A(x,y) = [\text{Grad}(m,n)]^P$ where P is a constant.

4.2.2 Mutual Information (MI)

MI can be determined from the following Eq. (14):

$$MI = \sum_{m=1}^L \sum_{n=1}^L P_{if}(m,n) \log_2 \frac{P_{if}(m,n)}{P_i(m)P_f(n)} \quad (14)$$

where, $P_{if}(m,n)$ represents the joint probability density distribution of the greyscale image in i and f . $P_i(m)$ and $P_f(n)$ represent the probability density distribution of the greyscale image in i and f , respectively. MI defines the sum of mutual information between each input image and the fused image. The greater MI shows that the fused image has greater information than the source images [29].

4.2.3 Feature Mutual Information (FMI)

FMI is the non-reference performance metric for fusion algorithm which calculates the mutual information of the image features, like edges and gradients [30, 31]. FMI can be computed from the following Eq. (15):

$$FMI_f^{ij} = \frac{1}{N} \sum_{n=1}^N \frac{I_n(f,i)}{H_n(f) + H_n(i)} + \frac{I_n(f,j)}{H_n(f) + H_n(j)} \quad (15)$$

where, N represents the number of sliding windows, $H_n(f)$ is the entropy of n^{th} window in image f , $I_n(f,i)$ is the regional mutual information between the n^{th} window of image f and i . FMI_f^{ij} evaluates how much of feature information of the source images has been transferred into fused image. Higher value of FMI_f^{ij} shows better quality of fused image.

4.2.4 Fusion Similarity Metric (FSM) Q_T

Q_T is the fusion quality metric which measures similarity in terms of luminance, contrast and structure between the source and fused images. Q_T finds above similarities in the source and fused image block by block through a sliding window. Q_T can be calculated from Eq. (16):

$$Q_T = \frac{1}{|N|} \sum_{n \in N} \zeta^{M,N} Q(i, f|n) + (1 - \zeta^{M,N} Q(j, f|n)) \quad (16)$$

where, i is the source image, f is the fused image and similarity map $\zeta^{M,N}$ is defined as in Eq. (17):

$$\zeta^{M,N} = \begin{cases} 0, & \frac{\varphi_{if}}{\varphi_{if} + \varphi_{jf}} < 0 \\ \frac{\varphi_{if}}{\varphi_{if} + \varphi_{jf}}, & 0 \leq \frac{\varphi_{if}}{\varphi_{if} + \varphi_{jf}} \leq 1 \\ 1, & \frac{\varphi_{if}}{\varphi_{if} + \varphi_{jf}} > 1 \end{cases} \quad (17)$$

where φ_{if} is defined as per Eq. (18):

$$\varphi_{if} = \frac{1}{K-1} \sum_{k=1}^K (i_k - \bar{i}) \cdot (f_k - \bar{f}) \quad (18)$$

4.2.5 Results Discussion

Proposed fusion scheme for multi-focus image fusion is compared with the state-of-the-art latest fusion algorithms which include DSIFT, DWT, DCT, GFF, QMIF, DCT-EoL and TFS. Four types of quality assessment metrics as described earlier were used to compare the effectiveness of our proposed fusion scheme.

Tab. 1 shows the value of SSS which is the measure of the edge information transferred in the fused image from source images. The overall SSS index of the proposed

method is better than the other techniques, for most of the source images. However, for "Pepsi", "Joy" and "Toy" source images, our results for SSS are approximately equal to TSL method. Our proposed method acquires the edges of source images for determining the initial fusion map. The edge information in near and far focused 'Pepsi' images is not well defined, hence, in-terms of SSS, our proposed method gives the results approximately the same as those of TSF. Similarly, for "Joy" and "Toy" source images, there is a minor mis-registration in both the source images and our proposed method specifically describes the effectiveness of the method on registered images, therefore, the SSS results of the proposed method for these images are approximately equal to the TSF. However, for the rest of the evaluation metrics *MI*, *FSM* and *FMI*, our proposed algorithm has produced significantly better results than the TSF for these three source images. Pertinent to mention here is that the SSS results of our proposed technique of above three images have a negligible margin as compared to TSF; hence it may be said that the maximum possible improvement of SSS for these three test images may have been achieved.

Tab. 2 has the second quantitative measure, i.e. *MI*. The *MI* directly measures the amount of common information between two source images. It can be observed from Tab. 2, that our proposed method has better *MI* for all the test images, except the "Temp.". For the multifocus image fusion, more information reflects the presence of higher magnitude of the gradients in the fused image. However, in Fig. 8a, it can be observed that few blocking artefacts are present at the boundary of near and far focused images, which increases the *MI* score for the QMIF method for most of the test images. Specifically, *MI* of 'Temp.', fused image, our results are approximately equal to QMIF.

Table 1 SSS for fused Multi-focus Images. The best results are shown in bold

Method	Ball.	Book	Clock	Disk	Doll	Flower	Girl	Jug	Lab	Leop.	N.P.	Path	Pepsi	Rose	Joy	Temp.	Toy	World
Prop.	0.83	0.725	0.719	0.739	0.745	0.696	0.720	0.783	0.749	0.828	0.682	0.694	0.784	0.720	0.746	0.768	0.692	0.775
DSIFT	0.828	0.719	0.711	0.729	0.732	0.694	0.718	0.779	0.744	0.826	0.674	0.689	0.763	0.710	0.742	0.762	0.676	0.774
DWT	0.825	0.699	0.693	0.704	0.714	0.654	0.701	0.751	0.724	0.822	0.577	0.594	0.77	0.683	0.719	0.677	0.657	0.723
DCT	0.828	0.719	0.711	0.732	0.729	0.678	0.715	0.778	0.746	0.826	0.673	0.679	0.785	0.709	0.741	0.746	0.689	0.773
GFF	0.828	0.717	0.711	0.725	0.697	0.694	0.715	0.757	0.738	0.827	0.664	0.69	0.777	0.707	0.743	0.758	0.696	0.772
QMIF	0.828	0.721	0.716	0.738	0.706	0.695	0.718	0.756	0.747	0.827	0.681	0.693	0.784	0.715	0.740	0.761	0.687	0.774
DCT-EoL	0.828	0.719	0.704	0.727	0.717	0.672	0.713	0.761	0.736	0.826	0.661	0.676	0.765	0.709	0.736	0.751	0.678	0.771
TSF	0.829	0.723	0.714	0.735	0.732	0.694	0.718	0.778	0.745	0.828	0.671	0.691	0.786	0.714	0.747	0.762	0.695	0.774

Table 2 MI for fused Multi-focus Images. The best results are shown in bold

Method	Ball.	Book	Clock	Disk	Doll	Flower	Girl	Jug	Lab	Leop.	N.P.	Path	Pepsi	Rose	Joy	Temp.	Toy	World
Prop.	11.21	9.31	8.67	8.31	8.21	8.08	8.81	8.61	8.81	10.94	6.53	9.09	9.03	7.75	8.29	7.36	8.4	9.49
DSIFT	11.18	8.98	8.38	7.9	7.93	8.01	8.69	8.56	8.50	10.89	6.36	8.96	8.22	7.54	8.25	7.15	7.5	9.40
DWT	10.87	8.29	7.83	7.37	7.42	6.64	8.05	7.98	8.21	10.52	5.15	7.62	8.35	6.71	7.78	6.11	6.7	8.15
DCT	11.18	9.2	8.6	8.27	7.81	7.85	8.73	8.52	8.80	10.90	6.45	8.92	8.67	7.64	8.22	7.26	8.14	9.46
GFF	11.13	8.6	7.88	7.06	6.77	7.32	8.03	7.77	7.91	10.93	4.48	8.08	7.37	6.58	7.81	5.78	6.88	8.77
QMIF	11.18	9.29	8.64	8.3	7.71	8.07	8.79	7.63	8.68	10.90	6.50	9.06	8.87	7.68	8.27	7.37	8.22	9.45
DCT-EoL	11.17	8.93	8.64	8.08	8.04	7.71	8.71	8.36	8.47	10.91	6.45	8.84	8.84	7.70	7.94	7.24	7.99	9.34
TSF	11.18	9.16	8.36	8.03	7.62	7.89	8.58	8.48	8.44	10.90	5.94	8.83	8.24	7.49	8.16	6.97	7.43	9.34

The *FMI* is a non-reference fusion metric which reflects the amount of feature information in the fused image. Tab. 3 shows that *FMI* of the proposed method is better than all other techniques, except the GFF for four test images. In case of *FMI*, GFF produces better results for "Ball.", "Leop.", "Toy" and "World" test images. From Fig. 9 (d) of the qualitative analysis, it can be observed that the

fused images produced by using GFF technique have some irrelevant information, which should not be present there. This additional information is used by *FMI* to calculate the feature information. Therefore, the *FMI* score of the GFF is better for a few of the test images, as compared to our proposed technique. The last evaluation metric for the quantitative comparison is the *FSM* and results are given

in Tab. 4. Similar to the previous evaluation metrics, overall performance of our proposed fusion scheme is better than the competing algorithms. However, for "Ball.",

"Book" and "Lab" images, our FSM results are approximately equal to the DSIFT, DCT-EoL and GFF, respectively.

Table 3 FMI for fused Multi-focus Images. The best results are shown in bold

Method	Ball.	Book	Clock	Disk	Doll	Flower	Girl	Jug	Lab	Leop.	N.P.	Path	Pepsi	Rose	Joy	Temp.	Toy	World
Prop.	0,924	0,912	0,926	0,926	0,916	0,892	0,887	0,88	0,920	0,849	0,772	0,825	0,925	0,901	0,882	0,825	0,919	0,894
DSIFT	0,923	0,907	0,924	0,912	0,911	0,890	0,886	0,87	0,919	0,848	0,770	0,82	0,924	0,896	0,881	0,82	0,914	0,893
DWT	0,923	0,906	0,921	0,91	0,911	0,877	0,886	0,868	0,919	0,848	0,771	0,819	0,924	0,896	0,880	0,819	0,911	0,889
DCT	0,923	0,907	0,924	0,911	0,911	0,886	0,886	0,869	0,918	0,848	0,762	0,817	0,924	0,893	0,881	0,821	0,916	0,894
GFF	0,996	0,873	0,895	0,883	0,85	0,892	0,856	0,864	0,915	0,896	0,733	0,745	0,88	0,843	0,881	0,588	0,930	0,907
QMIF	0,923	0,905	0,925	0,911	0,909	0,89	0,886	0,85	0,919	0,849	0,770	0,821	0,924	0,897	0,880	0,822	0,913	0,893
DCT-EoL	0,923	0,906	0,923	0,909	0,910	0,88	0,885	0,869	0,917	0,848	0,770	0,815	0,923	0,893	0,853	0,818	0,912	0,893
TSF	0,923	0,907	0,924	0,910	0,912	0,890	0,886	0,869	0,919	0,848	0,771	0,821	0,924	0,897	0,881	0,823	0,914	0,893

Table 4 FSM for fused Multi-focus Images. The best results are shown in bold

Method	Ball.	Book	Clock	Disk	Doll	Flower	Girl	Jug	Lab	Leop.	N.P.	Path	Pepsi	Rose	Joy	Temp.	Toy	World
Prop.	133,5	94,2	83,6	108,9	92,8	103,7	86,4	94,8	121,0	138,5	81,6	112,5	84,8	84,9	76,4	94,7	312,2	98,1
DSIFT	157,6	84,8	77,3	95,1	81,6	96,3	78,9	93,6	119,2	133,0	79,9	106,5	66,0	77,8	76,2	89,6	280,0	94,9
DWT	120,2	66,3	63,2	73,7	54,0	60,4	61,2	72,1	81,1	125,2	39,8	53,8	65,3	55,7	57,7	44,8	119,9	59,4
DCT	132,1	92,1	82,6	107,9	88,2	100,7	85,5	92,3	106,3	133,4	80,8	108,7	77,3	84,2	75,8	93,1	306,3	97,8
GFF	151,2	63,5	66,4	64,9	39,6	77,0	54,5	82,3	122,3	135,0	37,0	75,6	39,8	42,5	53,1	45,0	137,3	69,2
QMIF	129,2	92,8	80,2	104,2	80,7	102,4	85,8	76,6	102,9	134,4	78,1	111,7	70,2	81,3	75,4	90,0	163,5	96,3
DCT-EoL	155,1	95,1	82,1	100,3	80,1	94,6	81,9	84,4	97,1	133,2	81,1	104,2	71,2	81,9	59,2	92,8	289,9	96,8
TSF	141,8	86,9	75,5	92,3	67,9	96,85	78,9	89,4	95,4	133,3	73,9	102,7	64,1	76,6	72,5	83,7	288,9	92,2

5 CONCLUSION

In this paper, we proposed the SSGSM based multi-focus image fusion technique using multiple sized kernel. The source images are first pre-processed using the NMHE histogram equalization method and then gradients of these enhanced images are calculated using SSGSM. The basic focus map is determined from soft and hard decision maps acquired by employing multiple sized kernels. The final decision map is determined by applying morphological operations as a post processing step. The proposed algorithm is compared with the seven other state-of-the-art multi-focus image fusion algorithms for well known colour and grey multi-focus image dataset. It is concluded that the proposed algorithm has demonstrated significant improvement in qualitative results as compared to the rest of the algorithms. For quantitative comparison, four fusion metrics were used, because single fusion metric cannot show the effectiveness of a fusion scheme. For one fusion metric, our results are approximately equal to the best performing technique, but for the rest of the fusion metrics, our method performs way better than the same technique. Therefore, it is concluded that overall performance of our proposed technique in terms of quantitative measure was better than the rest of the techniques.

6 REFERENCES

[1] Li, H., Li, X., Yu, Z., & Mao, C. (2016). Multifocus image fusion by combining with mixed-order structure tensors and multiscale neighborhood. *Information Sciences*, 349, 25-49. <https://doi.org/10.1016/j.ins.2016.02.030>

[2] Xiao, J., Liu, T., Zhang, Y., Zou, B., Lei, J., & Li, Q. (2016). Multi-focus image fusion based on depth extraction with inhomogeneous diffusion equation. *Signal Processing*, 125, 171-186. <https://doi.org/10.1016/j.sigpro.2016.01.014>

[3] Amin-Naji, M. & Aghagolzadeh, A. (2018). Multi-focus image fusion in DCT domain using variance and energy of

Laplacian and correlation coefficient for visual sensor networks. *Journal of AI and Data Mining*, 6(2), 233-250.

[4] Farid, M. S., Mahmood, A., & Al-Maadeed, S. A. (2019). Multi-focus image fusion using content adaptive blurring. *Information fusion*, 45, 96-112. <https://doi.org/10.1016/j.inffus.2018.01.009>

[5] Scott, J. & Pusateri, M. A. (2010, October). Laplacian based image fusion. *2010 IEEE 39th Applied Imagery Pattern Recognition Workshop (AIPR)*, 1-7. <https://doi.org/10.1109/AIPR.2010.5759697>

[6] Song, Z., Jiang, H., & Li, S. (2017). A Novel Fusion Framework Based on Adaptive PCNN in NSCT Domain for Whole-Body PET and CT Images. *Computational and mathematical methods in medicine*. <https://doi.org/10.1155/2017/8407019>

[7] Singh, R. & Khare, A. (2014). Fusion of multimodal medical images using Daubechies complex wavelet transform-A multiresolution approach. *Information Fusion*, 19, 49-60. <https://doi.org/10.1016/j.inffus.2012.09.005>

[8] Cai, J., Cheng, Q., Peng, M., & Song, Y. (2017). Fusion of infrared and visible images based on nonsampled contourlet transform and sparse K-SVD dictionary learning. *Infrared Physics & Technology*, 82, 85-95. <https://doi.org/10.1016/j.infrared.2017.01.026>

[9] Gan, W., Wu, X., Wu, W., Yang, X., Ren, C., He, X., & Liu, K. (2015). Infrared and visible image fusion with the use of multi-scale edge-preserving decomposition and guided image filter. *Infrared Physics & Technology*, 72, 37-51. <https://doi.org/10.1016/j.infrared.2015.07.003>

[10] Liu, Y., Liu, S., & Wang, Z. (2015). A general framework for image fusion based on multi-scale transform and sparse representation. *Information Fusion*, 24, 147-164. <https://doi.org/10.1016/j.inffus.2014.09.004>

[11] Zhou, Z., Wang, B., Li, S., & Dong, M. (2016). Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters. *Information Fusion*, 30, 15-26. <https://doi.org/10.1016/j.inffus.2015.11.003>

[12] Li, X., Li, H., Yu, Z., & Kong, Y. (2015). Multifocus image fusion scheme based on the multiscale curvature in

- nonsampled contourlet transform domain. *Optical Engineering*, 54(7), 073115.
<https://doi.org/10.1117/1.OE.54.7.073115>
- [13] Bai, X., Zhang, Y., Zhou, F., & Xue, B. (2015). Quadtree-based multi-focus image fusion using a weighted focus-measure. *Information Fusion*, 22, 105-118.
<https://doi.org/10.1016/j.inffus.2014.05.003>
- [14] Zhang, X., Li, X., & Feng, Y. (2016). A new multifocus image fusion based on spectrum comparison. *Signal Processing*, 123, 127-142.
<https://doi.org/10.1016/j.sigpro.2016.01.006>
- [15] Aslantas, V. & Toprak, A. N. (2014). A pixel based multi-focus image fusion method. *Optics communications*, 332, 350-358. <https://doi.org/10.1016/j.optcom.2014.07.044>
- [16] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), 91-110.
<https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- [17] Liu, C., Yuen, J., & Torralba, A. (2010). Sift flow: Dense correspondence across scenes and its applications. *IEEE transactions on pattern analysis and machine intelligence*, 33(5), 978-994. <https://doi.org/10.1109/TPAMI.2010.147>
- [18] Liu, Y., Liu, S., & Wang, Z. (2015). Multi-focus image fusion with dense SIFT. *Information Fusion*, 23, 139-155.
<https://doi.org/10.1016/j.inffus.2014.05.004>
- [19] Mathew, J. J. & James, A. P. (2015). Spatial stimuli gradient sketch model. *IEEE Signal Processing Letters*, 22(9), 1336-1339. <https://doi.org/10.1109/LSP.2015.2404827>
- [20] Poddar, S., Tewary, S., Sharma, D., Karar, V., Ghosh, A., & Pal, S. K. (2013). Non-parametric modified histogram equalisation for contrast enhancement. *IET Image Processing*, 7(7), 641-652.
<https://doi.org/10.1049/iet-ipt.2012.0507>
- [21] Eskicioglu, A. M. & Fisher, P. S. (1995). Image quality measures and their performance. *IEEE Transactions on communications*, 43(12), 2959-2965.
<https://doi.org/10.1109/26.477498>
- [22] Nejati, M., Samavi, S., & Shirani, S. (2015). Multi-focus image fusion using dictionary-based sparse representation. *Information Fusion*, 25, 72-84.
<https://doi.org/10.1016/j.inffus.2014.10.004>
- [23] Liu, Y. & Wang, Z. (2013). Multi-focus image fusion based on wavelet transform and adaptive block. *Journal of image and graphics*, 18(11), 1435-1444.
- [24] Li, S., Kang, X., & Hu, J. (2013). Image fusion with guided filtering. *IEEE Transactions on Image processing*, 22(7), 2864-2875. <https://doi.org/10.1109/TIP.2013.2244222>
- [25] Haghghat, M. B. A., Aghagolzadeh, A., & Seyedarabi, H. (2011). Multi-focus image fusion for visual sensor networks in DCT domain. *Computers & Electrical Engineering*, 37(5), 789-797. <https://doi.org/10.1016/j.compeleceng.2011.04.016>
- [26] Ma, J., Zhou, Z., Wang, B., Miao, L., & Zong, H. (2019). Multi-focus image fusion using boosted random walks-based algorithm with two-scale focus maps. *Neurocomputing*, 335, 9-20. <https://doi.org/10.1016/j.neucom.2019.01.048>
- [27] Xue, W., Zhang, L., Mou, X., & Bovik, A. C. (2013). Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. *IEEE Transactions on Image Processing*, 23(2), 684-695.
<https://doi.org/10.1109/TIP.2013.2293423>
- [28] Yu, X., Chen, Q., Gu, G., Ren, J., & Sui, X. (2015). Perceived assessment metrics for visible and infrared color fused image quality without reference image. *Optical Review*, 22(1), 109-122.
<https://doi.org/10.1007/s10043-015-0058-9>
- [29] Qu, G., Zhang, D., & Yan, P. (2002). Information measure for performance of image fusion. *Electronics letters*, 38(7), 313-315. <https://doi.org/10.1049/el:20020212>
- [30] Haghghat, M. B. A., Aghagolzadeh, A., & Seyedarabi, H. (2011). A non-reference image fusion metric based on mutual information of image features. *Computers & Electrical Engineering*, 37(5), 744-756.
<https://doi.org/10.1016/j.compeleceng.2011.07.012>
- [31] Haghghat, M. & Razian, M. A. (2014, October). Fast-FMI: non-reference image fusion metric. *m2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT)*, 1-3.
<https://doi.org/10.1109/ICAICT.2014.7036000>

Contact information:

Muhammad MUZAMMIL, Lecturer
 (Corresponding author)
 International Islamic University, Islamabad, Pakistan,
 Room# 005, DEE, FET, IIUI, Sector H-10, Islamabad
 E-mail: m.muzammil@iiu.edu.pk

Imdad ALI, Senior Engineer
 NCP, Quaid-e-Azam University, Islamabad, Pakistan
 NCP, QAU, Islamabad, Pakistan
 E-mail: lmal507@gmail.com

Umer JAVED, Assistant Professor
 International Islamic University, Islamabad, Pakistan
 Room 201, DEE, FET, IIUI, Sector H-10, Islamabad
 E-mail: umer.javed@iiu.edu.pk

Muhammad AMIR, Professor
 International Islamic University, Islamabad, Pakistan
 DEE, FET, IIUI, Sector H-10, Islamabad
 E-mail: m.amir@iiu.edu.pk

Ihsan ULHAQ, Associate Professor
 International Islamic University, Islamabad, Pakistan
 DEE, FET, IIUI, Sector H-10, Islamabad
 E-mail: ihsanulhaq@iiu.edu.pk