IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 14, 2020, accepted December 2, 2020, date of publication December 7, 2020,
date of current version December 17, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3043048

A Novel De-Ghosting Image Fusion Technique for Multi-Exposure, Multi-Focus Images Using Guided Image Filtering

MUHAMMAD SOHAIB ROOMI[1], MUHAMMAD IMRAN[2], SYED ATTIQUE SHAH[1], AHMAD ALMOGREN[3], (Senior Member, IEEE), IHSAN ALI[4], AND MANSOUR ZUAIR[5]

[1]Department of Computer Science, Balochistan University of Information Technology, Engineering and Management Sciences, Quetta 87300, Pakistan
[2]Department of Electrical Engineering, Balochistan University of Information Technology, Engineering and Management Sciences, Quetta 87300, Pakistan
[3]Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11633, Saudi Arabia
[4]Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia
[5]Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

Corresponding authors: Syed Attique Shah (attique.shah@buitms.edu.pk) and Ihsan Ali (ihsanalichd@siswa.um.edu.my)

This work was supported in part by the University of Malaya under Postgraduate Research under Grant PG035-2016A, and in part by the Deanship of Scientific Research at King Saud University under Grant RGP-1437-035.

ABSTRACT In this paper, a novel de-ghosting image fusion technique is presented, which enhances the quality of low dynamic range images using multi-level exposures taken from the ordinary camera and also removes the ghosting artifact. In the proposed algorithm, first, the source images, taken under different exposure settings, are decomposed into base and detail layers using two-scale decomposition. The base and detail layers contain small and large-scale variation details of the source images, respectively. The Laplacian-of-Gaussian filter is applied to the source images to get the edge information. Afterward, the saliency map of the edges is computed. To remove the ghosting artifacts, a weight matrix is calculated by applying the median filter on the histogram equalized source images. The weight matrix is combined with the saliency map to generate more accurate weights. The separate weights for the base and detail layers are calculated using guided image filters. Finally, the base and detail layers' weights are fused with the source images to generate a vivid and enhanced image without any artifacts. The proposed technique is evaluated both qualitatively and quantitatively. The comparison of our technique in terms of Yang's Metric (Q_Y), Quality Mutual Information (Q_{MI}), Gradient-based Fusion Metric (Q_G) and Chen Blum's Metric (Q_{CB}) with other state-of-the-art techniques proves that the proposed technique outperforms existing techniques.

INDEX TERMS Deghosting, ghost effect, guided filter, high dynamic range imaging, multi-exposure image fusion, multi-focus image, weight map.

I. INTRODUCTION

The images captured by ordinary digital cameras do not contain the entire details of the real-world scenes [1]. This is due to the fact that the dynamic range of the real-world is large; whereas, the sensors deployed in ordinary cameras can capture only a tiny range of it [2]. The difference between the highest and lowest pixel values of an image is called its dynamic range. The resultant image captured by an ordinary camera loses the details because of the underexposed area of the image, which is dark due to low exposure, and some portion appears over bright due to high exposure. Therefore, digital images do not look as realistic as a human eye sees them, and so much detail is lost. There are two approaches to overcome this problem of significant difference between the High Dynamic Range (HDR) of the real-world and the Low Dynamic Range (LDR) of digital images, namely: the hardware-based approach and the software-based approach. In the hardware-based approach, cameras are equipped with sensors having HDR imaging capabilities. However, the state-of-the-art CMOS or InGaAs sensors that have the capability of capturing the HDR of the real-world [3], [4] are not affordable for everyday users. The software-based approach, on the other hand, is way less expensive and thus can be considered a more practical solution.

The associate editor coordinating the review of this manuscript and approving it for publication was Inês Domingues.

In the software-based approach, multiple images that are taken by ordinary cameras under different exposure settings are fused together to produce an HDR image that looks more realistic and contains more details [5]. There are two image fusion methods: tone mapping (TM) and multi-exposure image fusion (MEF). The TM method consists of two essential steps: HDR reconstruction and tone mapping. As first step, multiple LDR images of the same scene are combined to obtain an HDR image [6], [7]. However, the obtained image cannot be displayed on an ordinary camera or any LDR device directly due to its high fidelity. Therefore, tone mapping is applied to the HDR image to make it suitable for display. On the contrary, the MEF technique is simple and does not require any processing before displaying, which makes it more suitable [8], [9]. We have implemented the MEF technique in which multiple exposures of images are taken under different levels of brightness and fuse the informative part of the image into a single resultant image.

Although MEF techniques are considered more reliable and efficient than their counterpart tone mapping techniques, they have their specific limitations. In the case of static scenes, i.e., when there are no moving objects in the LDR images, MEF techniques work fine. However, MEF techniques results are not promising if the LDR images are taken under different exposure settings that may contain moving objects, called ghosting artifacts. In such cases, the resultant HDR images contain the shadows of ghosting artifacts, hence the ghosting artifact is removed in this paper. In addition to the above limitations, there are some other issues as well. For example, in practice, due to ripples or handheld camera, small-displacement also occurs. However, this small-displacement of camera movement could be tackled either by deploying a tripod or implementing some registration methods [10]–[13], and similarly for object motion, many available methods require pixel and patch level execution [14]–[18]. Therefore, in this research work, the camera movement is not considered.

The major contributions of this paper can be summarized as follows:

1) A single comprehensive method is developed which works on focused, flashed and multi-exposure images and also removes ghosting artifacts from the fused image.
2) Reducing the computational complexity in image fusion process with the help of guided image filtering.
3) A novel method of initial weight construction using Laplacian-of-Gaussian and a weight map with maximum pixel among all images is devised.
4) Resulting images are refined using the optimal radius and the regularization parameter of the guided filter.
5) The proposed technique is evaluated through comparison with various other state-of-the-art techniques. The results show finer performance superiority of the proposed technique.

The rest of the paper is structures as follows. In Section II, we review the existing multi-exposure image fusion, tone mapping and de-ghosting methods. Section III describes the proposed fusion method in detail. In Section IV, we compare our experimental results with state-of-the-art methods and conduct further discussions. In Section V, we conclude the paper.

II. RELATED WORK

An important aim of digital photography is to reproduce the natural scene with good contrast, vibrant color and rich imagery. Nevertheless, the captured images are sometimes under-exposed or over-exposed due to poor lighting environments and the restricted dynamic range of imaging equipment. The captured LDR images degenerate the performance of numerous computer vision and image analyzing techniques/algorithms. Thus LDR to HDR enhancement is an important step towards improving the efficiency of the captured images and making the detail of the image more rich and visible. An extensive amount of dedicated research studies on HDR imaging issues can be found in the existing literature. Due to the advantages of the HDR image a number of HDR imaging techniques have been proposed by [19]–[24].

Nayar and Mitsunaga [11] proposed the method to produce HDR image from different exposure after considering the global motion of camera to register images. It uses the Computed Response Function (CRF) to fuse multiple source images of different exposures into single HDR radiance image. Although, it is a good option for the infrared detector but in this technique for each pixel, an error function is required to be defined which increases the computation complexity. Ward [10] proposed the alignment method which uses percentile threshold bitmaps to speed up image operations and prevent problems of different levels of exposure used in photography with HDR. Cost of this method is linear in terms of the number of pixels and is completely independent of the total translation. Global strategies are therefore very effective, but they are unreliable in the presence of independently moving objects while zooming or tiling. Chen *et al.* [13] addresses the transformation of multi-exposure images taken by LDR device, using key points (or feature points) those are obtained by applying SIFT technique on source images. Content of image does not affect its performance and also works well for the under-and over-exposed images. However this method fails when camera motion is not centered and produce an artifact in output image. Reinhard *et al.* [25] presented a technique to develop an operator of tone reproduction, with the help of local contrast measurement of the multi-exposure images and constructing the the density of luminance for HDR image. Their technique provides an effective way to compress the dynamic range and also reduce halo artifact simultaneously. Nevertheless, its quality is restricted by the circular surrounding. Kuang *et al.* [26] introduced a HDR image rendering model named as iCAM06 that is based on the TM operator and worked with the strained color image. Their architecture accounts for rod in low light conditions, enabling accurate simulation of the HDR scene for the user view. However, this technique has difficulties for high level

luminance predication and loss the detail of spatial acuity at darker area. Shan [19] developed the TM operator which performs the local linear modification over all source images using a small overlapping window. This technique helps to synthesize the resultant image from LDR images. The typical problem with some TM methods such as this is halo artifact because of contrast reversals.

MEF techniques provide an alternate way of producing informative and perceptually appealing HDR images, which directly obtain the fuse image for LDR devices. MEF method using block based approach by Goshtasby [20] divides the image into non-overlapping blocks and select the blocks that have maximum information from all blocks also known as region based MEF technique. Optimal block size is define by the gradient-ascent algorithm to get the highly informative fuse image. If the color information within highlight area is comparatively high, it will preserve the scene information. However, if the block size is not sufficiently small, this technique may produce artifacts on the scene boundaries. Wang and Zhang [27] presented a novel multi-exposure image fusion method which works on patch segmentation. In this method they have used super-pixel segmentation then it is decomposed into three components and fused independently that is based on the human vision. Finally, the guided filter is used for optimized results. Ma *et al.* [28] introduced the MEF method which works with low-resolution of source images. They have implemented the fully conventional network and then through guided filter done the up-sampling and get the fused image after applying the weighted average. This method works for static scenes. Zhu [29] represented the novel MEF method, which is based on multi-modality, and image cartoon texture decomposition is used to preserve the structure of every source image. This technique preserves the details of the source image in the final resultant image. Mertens *et al.* [9] presented a persuasive weight based multi-exposure image fusion method. Which works on pixel level by calculating and combining the 3 superiority measures including saturation, contrast and well exposedness of each exposure image.It is computationally efficient and blending is reliable in preventing seams as it incorporate objects features. But the performance of these types of techniques could be unsatisfactory if the decomposition level is too small or too large. To overcome this issue, the proposed method uses two-scale image decomposition technique. Both the Song *et al.* [21] and Shen *et al.* [22] separately presented their probabilistic model-based multi-exposure image fusion methods. However method [22] works with two quality measures contrast and color-consistency and [21] calculates the level of image luminance then the images are embedded by gradient. These techniques are suitable for preserving the details, when an HDR image has a very large contrast ratio. But, because of multiple iterations these techniques may over smooth the resultant image. Gu *et al.* [23] fused gradient field which is obtained from the tensor structure of source based images on multi-dimensional Riemannian geometry then for modified gradient field iteration take twice mean filtering and for multi scale non-linearly compressing. No human interaction is needed in the execution to reproduce the results. Multi-scale decomposition used in technique slow down its performance, to overcome this issue we have use two scale decomposition. Ocampo-Blandon and Gousseau [30] used the Poisson editing technique to handle the saturated parts of multi-exposure images after that use the reference image by applying many patches for fusion and finally to emphasize stacking, non-local method implements. However, this technique seems suitable when there is less difference in saturated parts between multi-exposure images. Although it may produce artifacts on the edges of the fused image. Cai *et al.*'s [24] represented a different method for the image fusion of a single image by using the image contrast enhancement function which made the results that can be compared with multi-exposure image fusion. But still this method have some limitations for the darkest and brightest area of image and difficult to manage the large data-set. Also there is no consideration of moving objects.

In most current MEF methods, the basic assumption is that during various captures of multi-exposures the scene is static. But when fusing the images captured in dynamic scenes which involves moving objects, the above mentioned methods can produce ghosting artifacts. Various solutions are proposed such as Sidibe *et al.* [31] introduced one of the pixel level image fusion method using pixel order relation that helps to detect the moving objects with high sensitivity. If there is no motion object in pixel then the value of pixel intensity will be same as in exposure. After getting the moving objects it uses the static tool to get artifact free image and also work for the small background motion. Zhang and Cham [15] detected the ghost region by using gradient based approach. The two quality measures are introduced: Consistency and visibility, which is based on the change in the gradient of different exposures. This technique has lower computational complexity and also work for flash and non flash images. It may produce unsatisfactory results when the source images contain focused or un-focused images. Pece and Kautz [18] proposed a method for removing ghosting artifact by using motion-region based technique called BMD (Bitmap Movement Detection). They first extract the contrast, well-exposedness and saturation for each source image then detect the moving object by applying median bitmap. That help to impose relation from each exposure image. However it may produce some unrealistic effects on the fused image as it ignore image structural information. An *et al.* [32] modified the weights of [9] Mertens *et al.* for presenting the patch based correlation which converts the formulation of weight and represent as a motion pointer using photo-metric relation. The technique produce good texture detail and enhanced colored fused image. However this technique require more computation and also produce artifact on the shadow. Li and Kang [16] presented the method which collects three image features: brightness, local-contrast and color-dissimilarity to construct the weight map and enhance weight maps by using recursive filters. To avoid moving

object in fused image this technique use motion recognition function, which is accomplished by selecting the image as reference by using median filter and getting the histogram of remaining of static one. But it can not preserve the details when the source images severally have the brightest region.

Vanmali *et al.* [33] proposed MEF method for dynamic scene used four stages e.g weight map creation of source images, detection of moving object, than weight map adjustment and finally multi-exposure fusion via changed weights map. Method removes the ghosting artifact but unable to preserve the texture detail of darker area of image and also lost the color. To preserve the details of texture and natural color, proposed paper used the saliency map at each pixel level. Mertens *et al.* [9] came up with a method that works with contrast and saturation technique to fuse multiple exposure image sequence. This method produce good fusion efficiency also for the flash images but it cannot handle object boundaries well and some area of darker and brighter image lost the details. This issue is solved in proposed method by Vonikakis *et al.* [34], which works on illumination estimation based method in which well exposedness is calculated by a filtering method for illumination estimation. Then to create weight map, calculated estimates are combined with fuzzy membership functions in resultant well exposed fused image is generated. Ma *et al.* [14] presented the structural based patch decomposition method, which is disassembled in to three separate components, signal structure, mean intensity and signal strength, where each patch is fused independently and desired patches are placed in fused image. This method have very low computational cost as it also not required post processing steps. This method removes the ghosting artifact, however lost the details in darker area and produce blurred result. Lee *et al.* [35] worked with adaptive weight, relative pixel value and global gradient that calculate weight map for each exposure image and then obtain the compared intensity between the global gradient and source image. This method preserve the low luminance details better but often create unwanted visual artifacts. Paul *et al.* [36] proposed MEF method used gradient domain for color images, that is based on the luminance of source image using higher gradient magnitude at every intensity location and then via Haar wavelet technique obtain the fuse image. However, this method does not work with flash images. A proposed method by Bavirisetti *et al.* [37] covers this issue by combining information of image by multi scale image decomposition and perform visual saliency detection on it than finally construct weight map pixel by pixel on each scale. With the excessive brightness in fused image the color saturation is normalized.

There are many existing techniques to overcome the issue of LDR images in which some of the above methods have used a block-based technique. That divides the image into non-overlapping blocks and selects the maximum information from all blocks. Such techniques will preserve the scene information within the highlighted area and perform faster. However, if the block size is not optimal, it will produce artifacts on the scene boundaries. Some of the pixel level techniques used a weighted average by using different filters and optimizing the weight for a better quality image. The resultant fused images in such methods are vivid and reliable as it incorporates object features. But the performance of such techniques is not satisfactory because of the high or low level of decomposition. We have used the pixel-level technique in this paper, and to overcome the issue of decomposition, the two-level decomposition method is implemented and with the help of a guided filter, it produces optimal results on time.

III. PROPOSED METHOD

We propose a solution for multi-exposure and multi-focus images. The source images, taken under different exposure-settings or with diverse focuses, are first decomposed into base and detailed layers. Then, weights, called saliency map, for highly-structured and smooth areas are calculated using the Laplacian-of-Gaussian (LoG) filter. In the next step, which is especially well-suited for de-ghosting, the color dissimilarity features are estimated. The saliency map is refined using the guided image filter after combining it with the color dissimilarity features to get weight maps associated with the base and detailed layers. Finally, the base and detailed layers along with their respective weight maps, are fused together to generate a more vivid, enhanced, and detailed image without any ghosting artifacts. The proposed technique is comprehensively elaborated in the sequel.

A. OVERVIEW OF GUIDED IMAGE FILTER

In previous years, edge-preserving filters [38], [39] tended to be a most promising research topic in the image fusion field. There has been a number of edge-preserving filters implemented in image fusion methods such as conventional weighted least squares [39], bilateral filter [40], [41] and joint bilateral filter [42] in the early research, however, these filters may go through the consequences of ''gradient reversal'' artifact which occurs when a pixel has similar pixels around it, and also, a weighted average of Gaussian can be unstable. A Guided Filter (GF) has been proposed by He *et al.* [38], which significantly improved computing time. The GF has a fast and non-approximate linear time algorithm, whose computational complexity is independent of the filtering kernel size [38]. The GF can be utilized in many other applications such as image de-hazing [43], de-noising [44], removing snow/rain from images [45], soft matting, image enhancement and so on.

B. DE-GHOSTING IMAGE FUSION TECHNIQUE FOR MULTI-EXPOSURE, MULTI-FOCUS IMAGES USING GUIDED IMAGE FILTER

I_c^q represents the sequence of source images taken under different exposure settings and with diverse focuses; where, $q \in [1, n]$ and $c \in \{r, g, b\}$ represents, the number of source images and channels, respectively. The steps of the proposed method are elaborated in the following sub-sections.

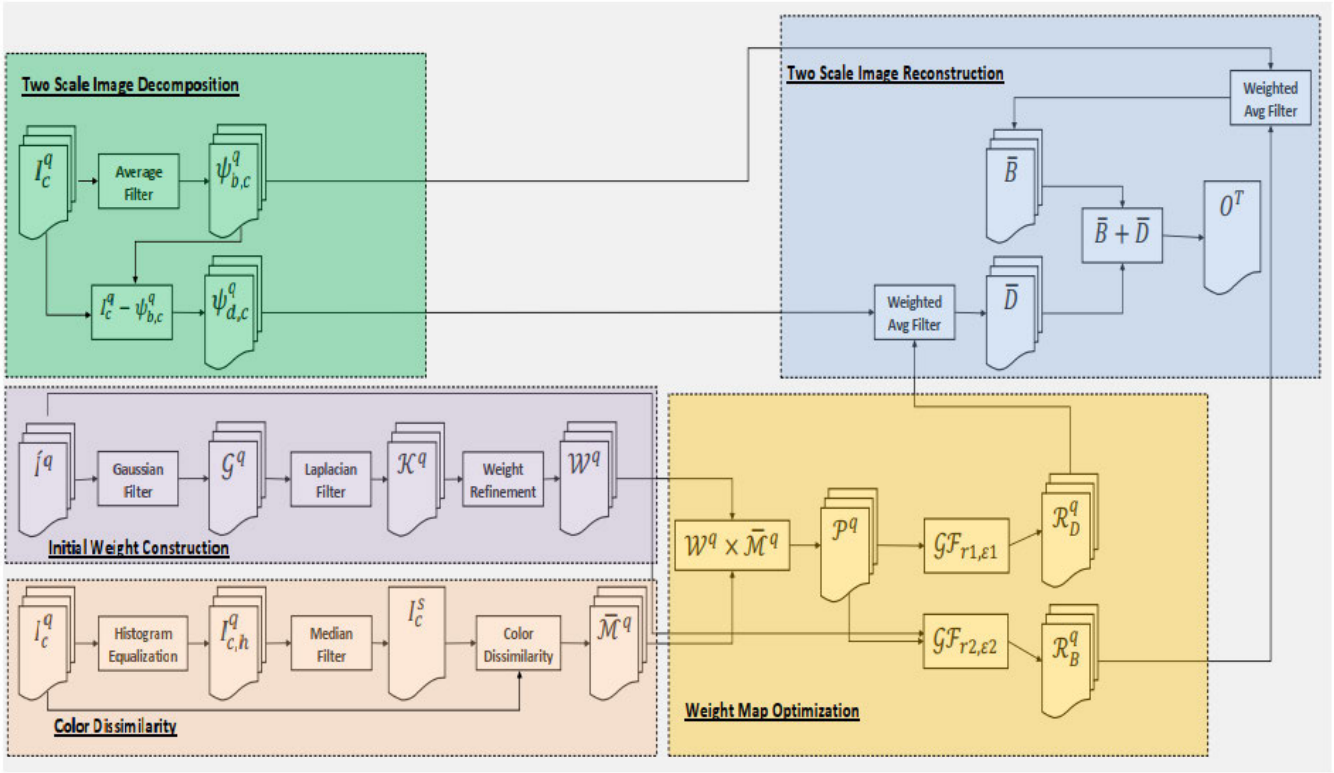


FIGURE 1. An overview of the proposed method.

Algorithm 1 Two Scale Image Decomposition

Input : I_c^q; Source Images.
Output: $\psi_{b,c}^q$; Base Layers, $\psi_{d,c}^q$; Detail Layers

$G = ToGrayScale(I_c^q)$
$se =$ AverageFilter(31 x 31) // conventionally set as 31
$N = size(I_c^q, 3)$ // getting the number of images
for $i \leftarrow 0$ **to** N **do**
  $\psi_{b,c}^i =$ meanfilter (G^i, Se) // Mean Filter Applied and got Base Layer
  $\psi_{d,c}^i = G^i - \psi_{b,c}^i$ // Detail Layer
end for
return $\psi_{b,c}^q, \psi_{d,c}^q$

1) TWO-SCALE IMAGE DECOMPOSITION

Two-scale image decomposition is used to get detail and base layer of source images, as shown in Algorithm 1. Base layer is extracted by applying an average filter on each source image which produces a smooth image. Detail layer which presents edges in the image are extracted after excluding the base layer from source images. To extract the information related to small-scale and large-scale variations in the sources images [46], called the base (ψ_b) and detailed (ψ_d) layers, an average filter is applied as follows,

$$\psi_{b,c}^q = \frac{1}{mn} \sum_{(i,j)\in S_{xy}} I_c^q(i,j), \tag{1}$$

$$\psi_{d,c}^q = I_c^q - \psi_{b,c}^q, \tag{2}$$

where, $q \in [1, n]$, $c \in \{r, g, b\}$, and S_{xy} represent the window size; that is, $m \times n$. In the proposed method, both m and n are chosen to be 31.

2) INITIAL WEIGHT MAP CONSTRUCTION

In this subsection, the weights for edges will be computed. Initially, Gaussian filter is applied to get a smooth image so that we can avoid artifacts. After that, a Laplacian filter is applied to the smooth image to get edges. As shown in Algorithm 2 edges from source images are taken as input and set weight maps as output. Remmet function normalizes it into zero to one range. Each pixel value of the image is compared with the relative pixel of other images and the maximum pixel value is set as 1 (weighted pixel, we are interested in) and the rest are set as 0.

First, the Gaussian filter with kernel size of 5×5 is applied on the source images as it removes the noise and smoothes the images, as shown below,

$$\mathcal{G}^q = \grave{I}^q \circledast f(x, y), \quad q \in [1, n]. \tag{3}$$

Algorithm 2 Initial Weight Construction

Input : $\grave{I}^q$; Gray Scale Source Images
Output: $\mathcal{W}^q$; Initial Weight Map

for $i \leftarrow 0$ **to** N **do**
  $\mathcal{G}^i = GaussianFilter(\grave{I}^i)$
  $\mathcal{K}^i = LaplacianFilter(\mathcal{G}^i)$
end for
$S = \mathcal{K}^q / RepeatMatrix(sum(\mathcal{K}^q, 3)[1, 1, N]$
  `// Normalize into 0 - 1`
$Labels = N$
for $j \leftarrow 0$ **to** N **do**
  $mono = zeros(r, c)$; `// rows (r) and columns (c) of image`
  $mono(Labels == i) = 1$;
  $\mathcal{W}^i(x, y) = mono$;
end for
return $\mathcal{W}^q$

In the above equation, $\grave{I}^q$ is the gray-scale version of I_c^q, and $f(x, y)$ is a low-pass Gaussian filter, which is defined below,

$$f(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2+y^2}{\sigma^2}}, \tag{4}$$

where $x \in [-2, 2]$, $y \in [-2, 2]$, and $\sigma^2 = 11$ is standard deviation parameter for smoothing. Afterwards, the Laplacian filter $\mathcal{L}(x, y)$ of size 3×3 is applied on the smooth image $\mathcal{G}(x, y)$ as follows,

$$\mathcal{K}^q = \mathcal{G}^q \circledast \Delta^2 f, \quad q \in [1, n], \tag{5}$$

where, $\Delta^2 f$ is a symmetric 3 x 3 Laplacian Kernel. This matrix is used to localize the edges, which is defined as follow:

$$\Delta^2 f = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 1 \end{bmatrix} mn. \tag{6}$$

The estimated weights give us good characterization of the coarse level. However, the weight are further refined as follows.

$$\mathcal{W}^q(x, y) = \begin{cases} 1, & if\ \mathcal{K}^q(x, y) \geq \mathcal{E}(x, y), \\ 0, & otherwise. \end{cases} \tag{7}$$

where,

$$\mathcal{E}(x, y) = \max\{\mathcal{K}^1(x, y), \ldots., \mathcal{K}^n(x, y)\}, \quad 0 \leq x, y \leq M, N. \tag{8}$$

3) COLOR-DISSIMILARITY

The images captured at different exposures and under different focus settings may contain moving objects. As a result, when these images are fused to make a more vivid and detailed image, the resultant image contains shadows of these moving objects called ghosting artifacts. Therefore, in devising an image fusion techniques, these moving objects should be removed. A number of techniques [47]–[50] have been proposed to cater this problem. However, the existing techniques have some limitations, such as time-consumption and require the users' input to select the reference image. To overcome all these shortcomings, a method called color dissimilarity [16] is utilized in this paper.

Algorithm 3 Color Dissimilarity

Input : I_c^q; Colored Source Images
Output: $\bar{\mathcal{M}}^q$; Moving Object

for $i \leftarrow 0$ **to** N **do**
  $I_{c,h}^i = histeq(I_c^i)$ `// histogram equalization of each image`
end for
$I_c^s = median(I_{c,h}^q)$
for $i \leftarrow 0$ **to** N **do**
  $diff = I_{c,h}^i - I_c^s$
  $diff = diff + 10^{-25}$
  $sig = 0.1$
  $dir = exp(-1 * (diff - 0)^2 / sig^2)$
end for
for $i \leftarrow 0$ **to** N **do**
  $se1 = strel(disk, n1)$ `// Disk Shape Structure` $n1 = 3$
  $se2 = strel(disk, n2)$ `//` $n2 = 30$
  $dir(i) = dilate(dir(i), se1)$ `// dilatation`
  $\bar{\mathcal{M}}^i = erode(dir(i), se2)$ `// erosion`
end for
return $\bar{\mathcal{M}}^q$

As shown in Algorithm 3, through the median filter we get the reference image and acquire the moving object by subtracting it from histogram equalized images. This algorithm helps to remove ghosting artifacts. First, the histogram equalization is applied on the source images I_c^q to get the more enhanced images, $I_{c,h}^q$. Afterwards, to select the static background, a median filter is applied on the source images separately as follows,

$$I_c^s = median\{I_{c,h}^1, I_{c,h}^2, \ldots, I_{c,h}^n\}, \tag{9}$$

where, I_c^s represents the static background. The moving objects appear less often than the intended objects; therefore, a median filter extracts the static background efficiently. Whereas, the moving objects are extracted by calculating the color dissimilarity as follows,

$$\mathcal{M}_c^q = e^{\frac{I_c^q - I_c^s}{\epsilon}}, \quad q \in [1, n]. \tag{10}$$

Afterwards, the information of moving objects in different current channels of the source images are combined as shown below.

$$\grave{\mathcal{M}}^q = \mathcal{M}_1^q \times \mathcal{M}_2^q \times \mathcal{M}_3^q, \quad q \in [1, n]. \tag{11}$$

(a) Low (b) Medium (c) High

FIGURE 2. Mud house image sequence taken under different exposures.

The moving objects detected in 11 are further refined using the erosion and dilation operations as follows.

$$\bar{\mathcal{M}}^q = (\dot{\mathcal{M}}^q \oplus s_1) \ominus s_2, \tag{12}$$

where, $\oplus$ and $\ominus$ represent the dilation and erosion operations, respectively. The disk-like structure elements: s_1 and s_2 are chosen with different radii, discussed in more detail in Section IV. The resultant $\bar{\mathcal{M}}^q$ contains the information of moving objects, and the process does not require the user's input in selecting the reference image.

4) WEIGHT ESTIMATION AND REFINEMENT

In this subsection, the weights related to edges and moving objects are merged in way that both types of information remain intact. The weights calculated in (7) and (12) contain, respectively, information related to edges and moving objects.

In Algorithm 4, we have used a guided filter which helps to remove the artifacts from the boundaries of the image. It preserves the details of edges using the guidance image which is the input image.

Algorithm 4 Weight Map Optimization

Input : $\grave{I}^q, \bar{\mathcal{M}}^q, \mathcal{W}^q$
Output: $\mathcal{R}_B^q, \mathcal{R}_D^q$; Guide Optimized Base, Detail Layers

$\mathcal{P}^q = \bar{\mathcal{M}}^q \times \mathcal{W}^q$
for $i \leftarrow 0$ **to** N **do**
  $mean_{\grave{I}^i} = f_{mean}(\grave{I}^i)$
  $mean_{\mathcal{P}^i} = f_{mean}(\mathcal{P}^i)$
  $corr_{\grave{I}^i} = f_{mean}(\grave{I}^i \times \grave{I}^i)$
  $corr_{\grave{I}^i\mathcal{P}^i} = f_{mean}(\grave{I}^i \times \mathcal{P}^i)$
  $var_{\grave{I}^i} = corr_{\grave{I}^i} - mean_{\grave{I}^i} \times mean_{\grave{I}^i}$
  $cov_{\grave{I}^i\mathcal{P}^i} = corr_{\grave{I}^i\mathcal{P}^i} - mean_{\grave{I}^i} \times mean_{\mathcal{P}^i}$
  $a = cov_{\grave{I}^i\mathcal{P}^i}/(var_{\grave{I}^i} + \epsilon)$
  $b = mean_{\mathcal{P}^i} - a. \times mean_{\grave{I}^i}$
  $mean_a = f_{mean}(a)\ mean_b = f_{mean}(b)$
  $\mathcal{R}^i = mean_a \times \grave{I}^i + mean_b$
end for
return $\mathcal{R}_B^q, \mathcal{R}_D^q$
/* f_{mean} is mean filter, ϵ is regularization parameter and r is radius */

To keep this information intact, the weights are combined as follows:

$$\mathcal{P}^q = \bar{\mathcal{M}}^q \times \mathcal{W}^q, \quad q \in [1, n]. \tag{13}$$

The weight map $\mathcal{P}^q$ are noisy and not aligned with the edges that causes artifacts in the final fused image. Spatial consistency [51], which means if the color and brightness of two adjacent pixels are same then their weights should be same, is considered as a possible solution. A famous spatial consistency approach [52], in which two terms: energy function and smoothness term, are calculated. The energy function and the smoothness term contain the pixel saliences and edge aligned weights, respectively. The weight function is, then, further optimized using global minimization technique. However, the spatial consistency techniques are considered inefficient. Therefore, in this paper, a better and reliable method is used. To refine the weights, a guided filter is applied on each input image $\grave{I}^q$ and the corresponding weight map $\mathcal{P}^q$. Furthermore, the $\grave{I}^q$ works as a guidance image.

$$\mathcal{R}_B^q = \mathcal{GF}_{r1,\varepsilon 1}(\mathcal{P}^q, \grave{I}^q) \tag{14}$$

$$\mathcal{R}_D^q = \mathcal{GF}_{r2,\varepsilon 2}(\mathcal{P}^q, \grave{I}^q) \tag{15}$$

In above equations, the parameters: $r1$ and $r2$ denote the local window radii, $_{\varepsilon 1}$ and $_{\varepsilon 2}$ represent the regularization parameters of the guided filter. The parameters are selected to accordingly to calculate the weight maps for base and details layers. The incentive of the proposed weight construction technique is as follow. According to GF [38] if the local variance is smaller at the location (i, j) then the place of the pixel will be in flat area of guidance image and value of ak will be close to 0 and resultant filtering out will be equal to $\bar{P}_x$ which is average of adjacent input pixels. Instead of this if the local variance is larger at location (i, j) then it means pixel is at the edge area of the image and it will be close to 1. In both the case pixel weight of the brightness and color will have same weights. This is exactly according to the principle of spatial consistency. As it has been defined before that if the base layer seems spatial smooth then its corresponding weight must also be spatial smooth otherwise artificial edges will be revealed. In contrast to this sharp and edge, aligned weights are preferred for detail layers, which can lose its details if the weights are over smoothed. That is why to fuse the base layers filter size and blur degree both are kept large however for detail layer, small blur degree and small filter size are preferred.

5) TWO SCALE IMAGE RECONSTRUCTION

In this section we present the two steps for reconstruction of image. In first step, weighted averaging process fuse the base and detail layers of different input images by following equations,

$$\bar{B} = \sum_{n=1}^{N} W_n^B B_n \tag{16}$$

FIGURE 3. Left to Right; (a) result of Vanmali *et al.* [33], (b) result of Mertens *et al.* [9], (c) result of Vonikakis *et al.* [34], (d) result of Ma *et al.* [14], (e) result of Lee *et al.* [35], (f) result of Paul *et al.* [36], (g) result of Bavirisetti *et al.* [37], and (h) result of our proposed method.

$$\bar{D} = \sum_{n=1}^{N} W_n^D D_n \tag{17}$$

In second step both the fused layers of base and detail are combined to get the output image.

$$O^T = \bar{B} + \bar{D} \tag{18}$$

As shown in Algorithm 5, reconstruction of images is done by combining the base and detail layers with a weighted base and weighted detail layer. Then to acquire a fused image both layers are further added to preserve the details of images.

Algorithm 5 Image Reconstruction

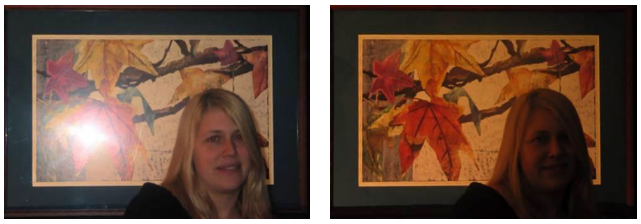
Input : $\mathcal{R}_B^q, \mathcal{R}_D^q, \psi_{b,c}^q, \psi_{d,c}^q$
Output: O^T
for $i \leftarrow 0$ **to** N **do**
  $\bar{B} = f_{mean}(\mathcal{R}_B^i * \psi_{b,c}^i)$
  $\bar{D} = f_{mean}(\mathcal{R}_D^i * \psi_{d,c}^i)$
end for
$O^T = \bar{B} + \bar{D}$
return O^T

IV. EXPERIMENTAL RESULTS

In this section, the experiments and results to validate the performance of the proposed scheme are presented. In addition to the objective and subjective evaluation of the presented technique, it is compared with prominent state-of-the-art techniques. Since the proposed technique provides a solution for both static and dynamic scenes, therefore, source images [24], [28], [53], [54] and the techniques for comparison [16], [55]–[58] are chosen accordingly. We compare our method with MEF and ghosting artifact removing techniques that are specifically presented for dynamic or static scenes.

A. EXPERIMENTAL SETUP

To evaluate the performance of the proposed method, experiments are conducted on different 20 image data-sets [53]. The image data-set includes sequences of images taken under different exposure setting, lighting conditions, focuses; few



(a) With Flash-Light (b) Without Flash-Light

FIGURE 4. Flash Image Sequence.

of them are static and others are dynamic. Among 20 image data-sets, 09 images are static scenes with multiple exposures (≥ 3) in which both indoor and outdoor images are included. In order to evaluate the dynamic scene, 03 images data-sets are used under different level of exposures and with moving object. Further in a static category more 07 multi-focused and 01 flashed image data-sets are utilized.

The proposed method is compared with 7 existing methods i.e., [9], [14], [33]–[37] and has tried to overcome the limitations in the existing methods both subjectively and objectively. In [33], the method removes the ghosting artifact but was unable to preserve the texture detail of the darker area of the image and unable to produce vivid color. To preserve the texture details and natural color, our proposed method constructed the saliency map at each pixel level. The method proposed in [9] produces good fusion efficiency also for the flash images but it cannot handle object boundaries, which is covered in our method by weight optimization. In [34], the authors suggested reducing the pyramid layer, in resultant this can produce halos, and thus we propose to refine the weight maps using the weighted average filter. In [35], the suggested method preserves the low luminance details effectively but often create unwanted visual artifacts. The method in [36] is based on the luminance of the source image using a higher gradient magnitude at every intensity location. It works well for focus images but for multi-exposure images, results are unsatisfactory in darker areas which can be seen in Figure 3f. The method in [37] works with multi-scale image decomposition and perform visual saliency detection on it. Usually, it smooths the high-frequency information of the image due to which it cannot detect the defocus or focus areas of the image. This can be further examined with Figure 11g, that shows another result of the same method having detection errors at the edge area of the image. Although guided image filter can improve the quality of the resultant image which avoids the block effect produce on the edge areas. The proposed technique works for the dynamic scene, multi-exposure static scene, multi-focus, and flash image.

B. SUBJECTIVE ASSESSMENT

As experiments are performed on 20 data-sets includes 96 source images but due to space constraints, we consider one or two data set from each class for the subjective analysis. The proposed technique is compared with 7 states of the art methods. We start by showing the excellence of the experimental static scene to authenticate our results of the proposed method. Figure 2 shows the sequence of ''Mud-House'' in which some features that are visible in one exposure disappear in the others due to over- or underexposure. Therefore, the basic goal of composition is to preserve all features present in the exposure sequence and make them visible in one image. The experimental results of different methods on the ''Mud-House'' sequences are shown in Figure 3. Where Figure 3a is obtained by Venmali *et al.*'s method, which has an overall dark appearance inside the house due to that things are not visible. Figure 3b, 3e results of Merten's and S. Lee's preserve the detail to a great degree with better color saturation. Besides, comprehensive observation reveals shadow of the table due to brightness. With V. Vonikakis's and Durga's method as shown in Figure 3c, and 3g respectively, due to high brightness, factor details are lost in the middle of the fused image in Figure 3c and in all part of Durga's method Figure 3g. Due to bad color saturation in the result of Kede's technique as shown in Figure 3d, the visual quality of the fused image is degraded. Similarly, Figure 3f also work well with brighter area of image but loss the detail inside door because of darker area which is proposed by Sujoy *et al.* The result of our proposed method as shown in Figure 3h, preserve all detail and has good visual quality.

Figure 4 represents the image sequence of ''Flash'' data-set. A pair of images with flash-light appearance i.e., Figure 4a and without i.e., Figure 4b are shown. Experimental results of Flash data set is shown in Figure 5. As can be seen in Figure 5a Vanmali *et al.* method, the area of face loss the detail due to darker result. Figure 5b Mertens *et al.*, results are a better than Vanmali *et al.* but due to brightness on leaves behind the color got faded on the face region. In Figure 5c, due to brightness on the left side of image colors are faded. In Figure 5d, 5f, 5g the white spot of flash-light appeared, which rather should be removed in fused images. 5e method is applied equally on all the area of image but due to excessive brightness, natural colors quality is lost. 5h represents our proposed method that overcomes all the above short-comes in the fused images.

Figure 6 presents more results of the proposed method on the sample sequence which is gathered from three sources [24], [28], [54]. Figure 6(a) shows five image exposures with low, medium and high exposures, while Figure 6(b) shows the fused results of the exposures.

To validate the efficiency of our proposed method in dynamic exposures, we have used the ''Forest'' data set having 4 different exposures of dynamic scene with a moving man as shown in Figure 7. Figure 8 is the fused image sequence of ''Forest'' data set with moving man. The results of Mertens *et al.* Figure 8a exhibit good color saturation but comprehensive observation reveals ghosting artifact. In Figure 8d results of Kede *et al.* is almost free from ghosting artifact apparent as compared to Mertens *et al.*, however their results have dark appearance and also lack in color saturation. It can be seen in rest of the Figures i.e., 8b, 8c, 8e, 8g that

FIGURE 5. Left to Right; (a) result of Vanmali *et al.* [33], (b) result of Mertens *et al.* [9], (c) result of Vonikakis *et al.* [34], (d) result of Ma *et al.* [14], (e) result of Lee *et al.* [35], (f) result of Paul *et al.* [36], (g) result of Bavirisetti *et al.* [37], and (h) result of our proposed method.

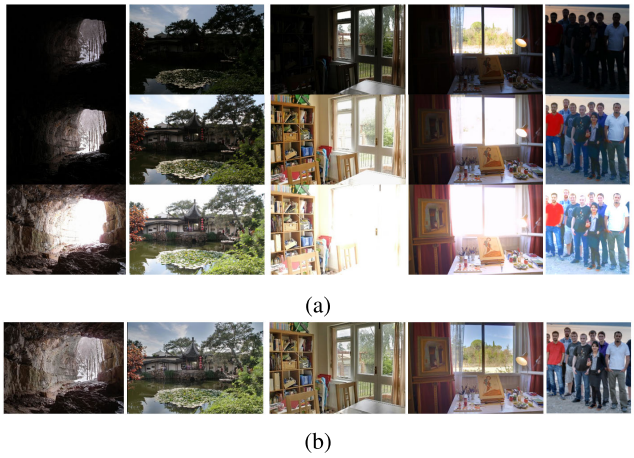


FIGURE 6. From left to right: (a) selected image exposures of the 'Cave', 'Chinese Garden', 'House', 'Studio' and 'Students' image sequences. (b) fused images produced by the proposed method.

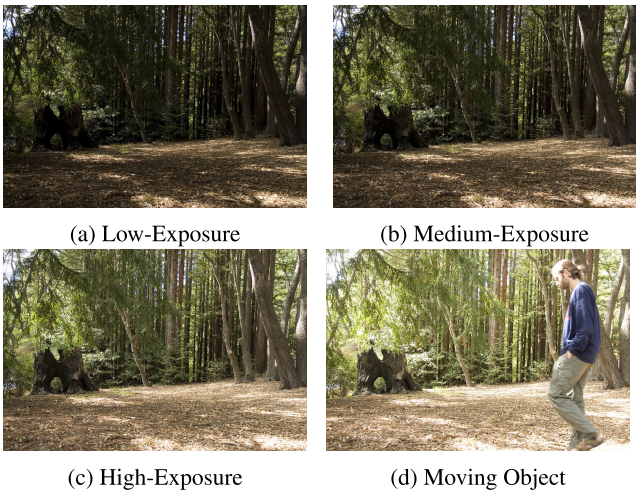


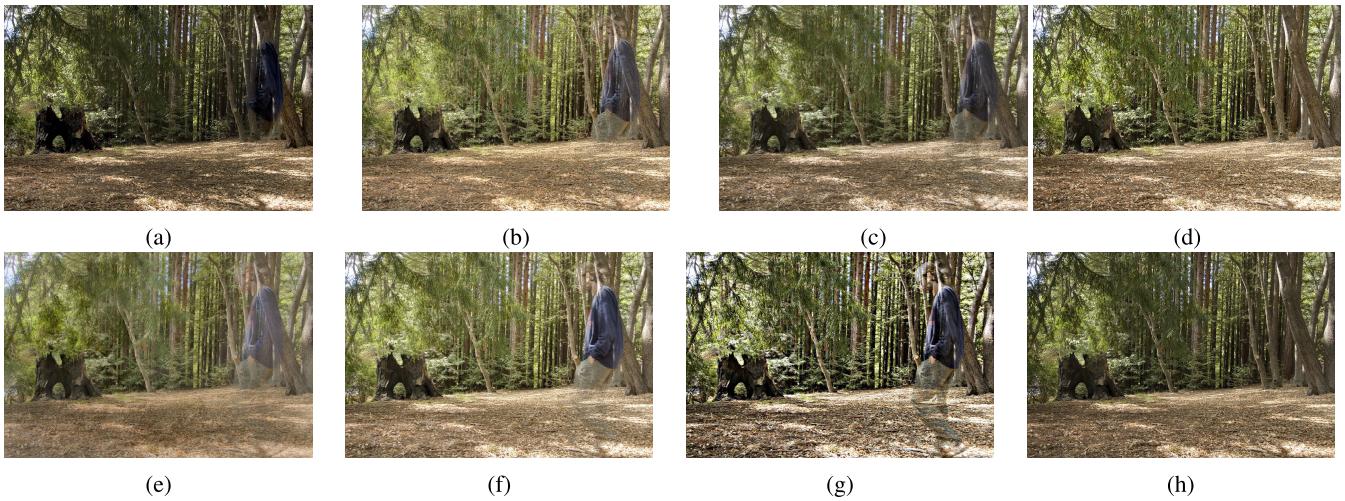
FIGURE 7. Forest Image Sequence.

the compared results suffers from severe ghosting artifacts. In contrast proposed method Figure. 8h shows pleasant visual results with more information as compared to the above results and is completely ghost-free.

Figure 10 presents the example of multi-focus color images data set named as "Children". To examine the performance of multi-focus, source images with a combination of different focus settings are used. To evaluate the performance of multi-focus, input images with a combination of different focus settings are used. Figure 11 is representing "Children" data set fused images obtained from multi-focus images of children sequence. In Vanmali *et al.*, Figure 11a portion of the image where girl standing is cleared, however the boy on the front has lost its details due to blurriness at shirt's collar of boy. Figure 11b of Tom Merten's has better results as compared to Vanmali et al's result, but the area of the statue is blurred. In Figure 11c, 11d the girl at the back with the statue is losing its details because that area is blurred. In Figure 11e, 11f the results are better to a great extent but the white area of the statue is not cleared. In Figure 11g the details have been lost due to excessive brightness which can be estimated from the boys's hair brightness. In our proposed method as shown in Figure 11h, the whole area at the front and back is highly focused and the visual quality is better.

C. OBJECTIVE ASSESSMENT

In order to assess the performance of fusion results quantitatively, fusion metrics or visual quality technique can be used. Mostly in image processing methods, resultant image can be justify by the visual inspection, which is the best way of quality assessment [59]. There is no single standard evaluation metric which shows the quality of resultant image and also



(a) (b) (c) (d)

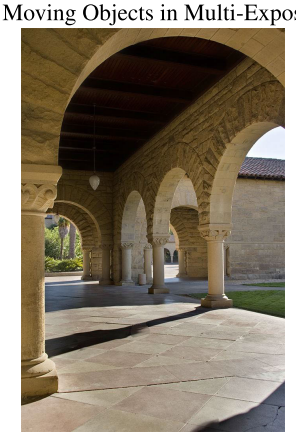
(e) (f) (g) (h)

FIGURE 8. Left to Right; (a) result of Vanmali *et al.* [33], (b) result of Mertens *et al.* [9], (c) result of Vonikakis *et al.* [34], (d) result of Ma *et al.* [14], (e) result of Lee *et al.* [35], (f) result of Paul *et al.* [36], (g) result of Bavirisetti *et al.* [37], and (h) result of our proposed method.

TABLE 1. Q_Y score of proposed method and 7 state of the art methods.

Data Set	[33] Vanmali	[9] Tom Merten	[34] Vonikakis	[14] Kede Ma	[35] S Lee	[36] Sujoy Paul	[37] Durga Prasad	Proposed
Mud-House	0.5567	0.5749	0.5376	0.4260	0.5834	0.5877	0.5810	**0.5977**
Children	0.9501	0.9150	0.8603	0.8836	0.8374	0.8825	0.8772	**0.9793**
Flash	0.7837	0.8738	0.8245	0.8604	0.6606	0.6789	0.6675	**0.8889**
Forest	0.5570	0.4933	0.5231	0.6226	0.4091	0.4915	0.5437	**0.6847**

(a) Moving Objects in Multi-Exposures



(b) Fused image

FIGURE 9. Arch image sequence for Multi-Exposure De-Ghosting.

(a) Front Focused (b) Back Focused

FIGURE 10. Children image sequence of multi-focus.

no standard choosing method for selecting fusion metric for evaluation. Therefore, we have applied different recent and traditional comprehensive evaluation metrics such as, structure based evaluation technique Q_Y [55], mutual information theory based assessment method Q_{MI} [56], visual image evaluation method which is inspired by human perception to assess the human visualization results of fused image Q_G [58] and Q_{CB} [57], [60].

1) Q_Y

Q_Y method defined by Yang *et al.* [55] works on the similarity of structure for fusion evaluation. Following equation define Q_Y as,

$$Q_Y = \begin{cases} \lambda_w SSIM(A_w, F_w) + (1-\lambda_w) SSIM(B_w, F_w), \\ \quad if\ SSIM(A_w, B_w|w) \geq 0.75 \\ max\ SSIM(A_w, F_w), SSIM(B_w, F_w), \\ \quad if\ SSIM(A_w, B_w|w) < 0.75 \end{cases} \tag{19}$$

FIGURE 11. Left to Right; (a) result of Vanmali *et al.* [33], (b) result of Mertens *et al.* [9], (c) result of Vonikakis *et al.* [34], (d) result of Ma *et al.* [14], (e) result of Lee *et al.* [35], (f) result of Paul *et al.* [36], (g) result of Bavirisetti *et al.* [37], and (h) result of our proposed method.

TABLE 2. Q_{CB} score of proposed method and 7 state-of-the-art methods.

Data Set	[33] Vanmali	[9] Tom Merten	[34] Vonikakis	[14] Kede Ma	[35] S Lee	[36] Sujoy Paul	[37] Durga Prasad	Proposed
Mud-House	0.4903	0.5390	0.4710	0.4267	0.4926	0.5401	0.5344	**0.5492**
Children	0.6408	0.6310	0.5958	0.5838	0.5638	0.5824	0.6323	**0.7730**
Flash	0.6895	0.6270	0.6364	0.6209	0.6355	0.6109	0.5562	**0.6932**
Forest	0.5801	0.4651	0.5099	**0.6035**	0.3906	0.4866	0.5961	0.5988

TABLE 3. Q_G score of proposed method and 7 state-of-the-art methods.

Data Set	[33] Vanmali	[9] Tom Merten	[34] Vonikakis	[14] Kede Ma	[35] S Lee	[36] Sujoy Paul	[37] Durga Prasad	Proposed
Mud-House	0.0516	0.1776	0.0489	0.0272	0.1731	0.0011	0.1678	**0.1838**
Children	0.775	0.7505	0.6801	0.7621	0.7689	0.5856	0.7602	**0.7751**
Flash	0.4432	0.6493	0.7583	0.7467	0.4145	0.6801	0.4492	**0.7606**
Forest	0.3207	0.3709	0.3626	0.3456	0.3322	0.3657	0.3066	**0.3721**

TABLE 4. Q_{MI} score of proposed method and 7 state-of-the-art methods.

Data Set	[33] Vanmali	[9] Tom Merten	[34] Vonikakis	[14] Kede Ma	[35] S Lee	[36] Sujoy Paul	[37] Durga Prasad	Proposed
Mud-House	0.4164	0.4194	0.3100	0.3956	0.3948	0.3902	0.4237	**0.4466**
Children	1.0156	0.8647	0.8017	0.8767	0.5641	0.882	0.7706	**1.0574**
Flash	0.8038	0.7934	0.7522	0.7731	0.5854	0.7646	0.5779	**0.8402**
Forest	0.3169	0.2784	0.3125	0.3083	0.2244	0.26865	0.3889	**0.4076**

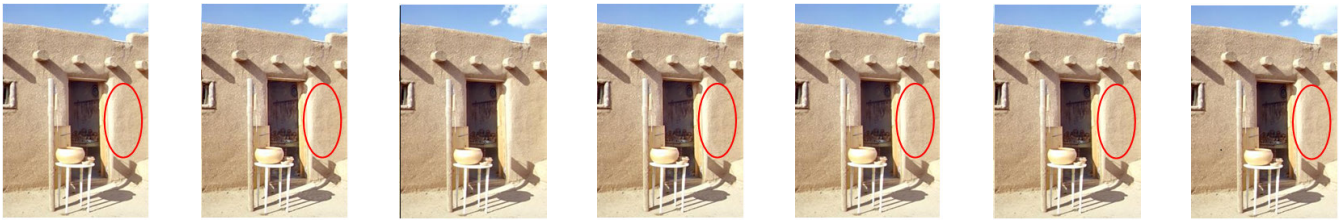
In the above equation input images are presented by A,B and fused image is denoted by F. However window size is set to 7×7, which is denoted by w and function $SSIM$ shows the structure similarity check.

$$\lambda_w = \frac{s(A_w)}{s(A_w) + s(B_w)} \tag{20}$$

To calculate the local weight λ_w presented in above equation in which variance of source images are represented by Aw and Bw from input image A and B in window w. This method preserve how well structural information of input images are gathered. Table 1 shows the Q_Y comparison of [9], [14], [33]–[37] with proposed method on 4 different source image sequences. Values with bold text are high in score and have better quality information.

2) Q_{MI}

Quality Mutual Information (Q_{MI}) technique works on information theory based metric. This technique have issue with



(a) r1:35, r2:1 (b) r1:40, r2:4 (c) r1:45, r2:7 (d) r1:50, r2:10 (e) r1:55, r2:13 (f) r1:60, r2:16 (g) r1:70, r2:22

FIGURE 12. Qualitative analysis of different parameters.

traditional MI metric [61] which is not stable and could make ambiguity measure for the source image with maximum entropy. This method is normalized by Turitsyna and Webb *et al.* [56] and defined as below,

$$Q_{MI} = 2\left[\frac{MI(A,F)}{H(A) + H(F)} + \frac{MI(B,F)}{H(B) + H(F)}\right] \quad (21)$$

The equation represents the input image A, B and output image F with their marginal entropy $H(A), H(B)$ and $H(F)$. Mutual Information (MI) between both F and A is measured through $MI(A, F)$ as mentioned below,

$$MI(A,F) = H(A) + H(F) - H(A,F) \quad (22)$$

Joint entropy between fused image F and input image A is denoted by $H(A, F)$ and $MI(B, F)$ and can be calculated same like $MI(A, F)$. This method inquire how efficiently the details of input image is saved in fused image. Table 4 indicate that the respective method performs well for the static and dynamic scene images, for multi-focus as well as flash images. It consistently leads to better output over other analyzed state of the art methods.

3) Q_{CB}

In this technique Chen and Blum [57] presents a contrast sensitivity function. Filter is applied on input and fused image after that for each image local contrast map is executed. Then relationship between source and fused image is described by preservation map. Finally overall quality is obtain by saliency map. Table 2 shows the execution results in which a larger value represents the better contrast. As per this metric, proposed method is considered as superior on other compared methods.

4) Q_G

This technique was presented by Xydeas and Petrovic [58] and the basic concept of the matrix is to preserve the maximum border of source images to fused image. This technique works on the sobel operator to calculate the orientation and strength of the gradient of each pixel at the resultant image. To evaluate the results, Table 3 shows the gathered statics. As per Q_G results, our proposed fusion method delivers the better performance to preserve the border details.

TABLE 5. Quantitative analysis of different parameters.

r1	r2	QMI	QG	QY	QCB
35	1	0.3675	**0.3022**	0.5701	0.5146
40	4	0.4357	0.252	0.5971	0.5394
45	7	**0.4466**	0.1838	**0.5977**	**0.5492**
50	10	0.4394	0.1775	0.5976	0.5474
55	13	0.4443	0.0832	0.5901	0.5448
60	16	0.4273	0.184	0.5973	0.5453
70	22	0.3973	0.2036	0.5971	0.5383

TABLE 6. Selected Parameters.

r1	$\epsilon 1$	r2	$\epsilon 2$
45	0.3	7	10^-6

TABLE 7. Consuming time of proposed method and 7 state-of-the-art methods on the image sequence of "Mud-House".

State-of-the-art Methods	Consuming Time
[33] Vanmali	4.23
[9] Tom Merten	3.31
[34] Vonikakis	4.53
[14] Kede Ma	4.02
[35] S Lee	3.33
[36] Sujoy Paul	3.49
[37] Durga Prasad	4.11
Proposed	2.93

D. ANALYSIS OF FREE PARAMETER

In this section, the results have been released by setting different parameters of radius r1 and r2, to select the best-fit parameter that will induce better-fused image. Each combination of the parameter is analyzed both quantitatively and qualitatively. For quantitatively, fusion quality performance is evaluated from the above-mentioned quality metrics. Table 5 shows the analyzed data of different parameters and bold numbers represent the best results. For qualitatively, Figure 12 shows that the brightness of the image is changed on every combination of parameters and selected parameters have better-equalized brightness in Figure 12c.

Parameters in Table 6 are induced for the guided filter.

E. EXECUTION TIME COMPARISON

After fusion quality, another important efficiency check for real application is computing evaluations. Table 7 Represents the comparison of average computational time with state-of-the-art methods on static image ''Mud-House''. All evaluations are tested on a CPU 2.3 GHz with 4GB RAM and using MATLAB R2017b.

V. CONCLUSION AND FUTURE WORK

In this paper, we have presented a novel multi-exposure method which is based upon a guided filter and color dissimilarity due to which it generates weight functions comprehensively and efficiently. We initially utilized a useful, simple and reliable filtering technique which is a guided image filter used for its edge-preserving and smoothness features. Secondly, we have combined histogram equalization and median filter for removing the moving objects in it. Final results after various experiments show that the proposed method is beneficial to preserve the details of source image, the fusion of multi-exposure images and for eliminating the moving objects. The proposed technique can be useful for various real-world applications such as machine vision for object detection and medical imaging. In future work, we will amplify its de-ghosting method for vanishing the moving objects in multiple locations and to secure more information by getting the desired area of an image. We also aim to make this method faster and easily applicable in mobile devices due to its low computational complexity.

REFERENCES

[1] S. Li, X. Kang, L. Fang, J. Hu, and H. Yin, ''Pixel-level image fusion: A survey of the state of the art,'' *Inf. Fusion*, vol. 33, pp. 100–112, Jan. 2017.

[2] S. W. Hasinoff, D. Sharlet, R. Geiss, A. Adams, J. T. Barron, F. Kainz, J. Chen, and M. Levoy, ''Burst photography for high dynamic range and low-light imaging on mobile cameras,'' *ACM Trans. Graph.*, vol. 35, no. 6, pp. 1–12, Nov. 2016.

[3] E. R. Fossum and D. B. Hondongwa, ''A review of the pinned photodiode for CCD and CMOS image sensors,'' *IEEE J. Electron Devices Soc.*, vol. 2, no. 3, pp. 33–43, May 2014.

[4] E. R. Fossum, ''Active pixel sensors: Are CCDs dinosaurs?'' in *Charge-Coupled Devices and Solid State Optical Sensors III*, vol. 1900, M. M. Blouke, Ed. Bellingham, WA, USA: SPIE, 1993, pp. 2–14.

[5] P. Ke, C. Jung, and Y. Fang, ''Perceptual multi-exposure image fusion with overall image quality index and local saturation,'' *Multimedia Syst.*, vol. 23, no. 2, pp. 239–250, Mar. 2017.

[6] P. E. Debevec and J. Malik, ''Recovering high dynamic range radiance maps from photographs,'' in *Proc. ACM SIGGRAPH Classes*, New York, NY, USA, 2008, pp. 31:1–31:10.

[7] S. Lee, G. H. An, and S.-J. Kang, ''Deep chain Hdri: Reconstructing a high dynamic range image from a single low dynamic range image,'' *IEEE Access*, vol. 6, pp. 49913–49924, 2018.

[8] K. Ma, Z. Duanmu, H. Yeganeh, and Z. Wang, ''Multi-exposure image fusion by optimizing a structural similarity index,'' *IEEE Trans. Comput. Imag.*, vol. 4, no. 1, pp. 60–72, Mar. 2018.

[9] T. Mertens, J. Kautz, and F. Van Reeth, ''Exposure fusion: A simple and practical alternative to high dynamic range photography,'' *Comput. Graph. Forum*, vol. 28, no. 1, pp. 161–171, Mar. 2009.

[10] G. Ward, ''Fast, robust image registration for compositing high dynamic range photographs from hand-held exposures,'' *J. Graphics Tools*, vol. 8, no. 2, pp. 329–344, 2005.

[11] T. Mitsunaga and S. K. Nayar, ''Radiometric self calibration,'' in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, Jun. 1999, pp. 374–380.

[12] K. Jacobs, C. Loscos, and G. Ward, ''Automatic high-dynamic range image generation for dynamic scenes,'' *IEEE Comput. Graph. Appl.*, vol. 28, no. 2, pp. 84–93, Mar. 2008.

[13] L. C. Chen, H. Y. Lin, S. K. Hung, M. S. Lee, W. Y. Chiou, Y. T. Shih, Y. C. Lo, Y. A. Lin, Y. H. Lin, L. W. Huang, E. S. Chung, and W. S. Chang, ''Applying a Web-based integrated radiation oncology information platform to enhance working efficiency and increase patient safety,'' *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 99, no. 2, p. E549, Oct. 2017.

[14] K. Ma, H. Li, H. Yong, Z. Wang, D. Meng, and L. Zhang, ''Robust multi-exposure image fusion: A structural patch decomposition approach,'' *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2519–2532, May 2017.

[15] W. Zhang and W.-K. Cham, ''Gradient-directed composition of multi-exposure images,'' in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 530–536.

[16] S. Li and X. Kang, ''Fast multi-exposure image fusion with median filter and recursive filter,'' *IEEE Trans. Consum. Electron.*, vol. 58, no. 2, pp. 626–632, May 2012.

[17] Q. Yan, Y. Zhu, and Y. Zhang, ''Robust artifact-free high dynamic range imaging of dynamic scenes,'' *Multimedia Tools Appl.*, vol. 78, no. 9, pp. 11487–11505, 2019.

[18] F. Pece and J. Kautz, ''Bitmap movement detection: HDR for dynamic scenes,'' in *Proc. Conf. Vis. Media Prod.*, vol. 10, no. 2, Nov. 2010, pp. 1–8.

[19] Q. Shan, J. Jia, and M. S. Brown, ''Globally optimized linear windowed tone mapping,'' *IEEE Trans. Vis. Comput. Graphics*, vol. 16, no. 4, pp. 663–675, Jul. 2010.

[20] A. A. Goshtasby, ''Fusion of multi-exposure images,'' *Image Vis. Comput.*, vol. 23, no. 6, pp. 611–618, Jun. 2005.

[21] M. Song, D. Tao, C. Chen, J. Bu, J. Luo, and C. Zhang, ''Probabilistic exposure fusion,'' *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 341–357, Jan. 2012.

[22] R. Shen, I. Cheng, J. Shi, and A. Basu, ''Generalized random walks for fusion of multi-exposure images,'' *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3634–3646, Dec. 2011.

[23] B. Gu, W. Li, J. Wong, M. Zhu, and M. Wang, ''Gradient field multi-exposure images fusion for high dynamic range image visualization,'' *J. Vis. Commun. Image Represent.*, vol. 23, no. 4, pp. 604–610, May 2012.

[24] J. Cai, S. Gu, and L. Zhang, ''Learning a deep single image contrast enhancer from multi-exposure images,'' *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 2049–2062, Apr. 2018.

[25] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, ''Photographic tone reproduction for digital images,'' in *Proc. 29th Annu. Conf. Comput. Graph. Interact. Techn. (SIGGRAPH)*, 2002, pp. 267–276.

[26] J. Kuang, G. M. Johnson, and M. D. Fairchild, ''ICAM06: A refined image appearance model for HDR image rendering,'' *J. Vis. Commun. Image Represent.*, vol. 18, no. 5, pp. 406–414, Oct. 2007.

[27] S. Wang and Y. Zhao, ''A novel patch-based multi-exposure image fusion using super-pixel segmentation,'' *IEEE Access*, vol. 8, pp. 39034–39045, 2020.

[28] K. Ma, Z. Duanmu, H. Zhu, Y. Fang, and Z. Wang, ''Deep guided learning for fast multi-exposure image fusion,'' *IEEE Trans. Image Process.*, vol. 29, pp. 2808–2819, 2020.

[29] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi, ''A novel multi-modality image fusion method based on image decomposition and sparse representation,'' *Inf. Sci.*, vol. 432, pp. 516–529, Mar. 2018.

[30] C. F. Ocampo-Blandon and Y. Gousseau, ''Non local exposure fusion,'' in *Proc. Prog. Pattern Recognit., Image Anal., Comput. Vis., Appl.*, Lima, Peru, Nov. 2016, pp. 484–492.

[31] D. Sidibe, W. Puech, and O. Strauss, ''Ghost detection and removal in high dynamic range images,'' in *Proc. 17th Eur. Signal Process. Conf.*, 2009, pp. 2240–2244.

[32] J. An, S. H. Lee, J. G. Kuk, and N. I. Cho, ''A multi-exposure image fusion algorithm without ghost effect,'' in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 1565–1568.

[33] A. V. Vanmali, S. G. Kelkar, and V. M. Gadre, ''Multi-exposure image fusion for dynamic scenes without ghost effect,'' in *Proc. 21st Nat. Conf. Commun. (NCC)*, Feb. 2015, pp. 1–6.

[34] V. Vonikakis, O. Bouzos, and I. Andreadis, ''Multi-exposure image fusion based on illumination estimation,'' in *Proc. Signal Image Process. Appl. Artif. Intell. Soft Comput.*, 2011, pp. 135–142.

[35] S.-H. Lee, J. S. Park, and N. I. Cho, ''A multi-exposure image fusion based on the adaptive weights reflecting the relative pixel intensity and global gradient,'' in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 1737–1741.

[36] S. Paul, I. S. Sevcenco, and P. Agathoklis, "Multi-exposure and multi-focus image fusion in gradient domain," *J. Circuits, Syst. Comput.*, vol. 25, no. 10, pp. 1–18, 2016.
[37] D. P. Bavirisetti, G. Xiao, J. Zhao, R. Dhuli, and G. Liu, "Multi-scale guided image and video fusion: A fast and efficient approach," *Circuits, Syst., Signal Process.*, vol. 38, no. 12, pp. 5576–5605, Dec. 2019.
[38] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.
[39] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," in *Proc. ACM SIGGRAPH Papers*, 2008, pp. 1–10.
[40] K. Kotwal and S. Chaudhuri, "Visualization of hyperspectral images using bilateral filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2308–2316, May 2010.
[41] F. Banterle, M. Corsini, P. Cignoni, and R. Scopigno, "A low-memory, straightforward and fast bilateral filter through subsampling in spatial domain," *Comput. Graph. Forum*, vol. 31, no. 1, pp. 19–32, Feb. 2012.
[42] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 664–672, Aug. 2004.
[43] C. Xiao and J. Gan, "Fast image dehazing using guided joint bilateral filter," *Vis. Comput.*, vol. 28, nos. 6–8, pp. 713–721, Jun. 2012.
[44] Q. Zhang and X. Li, "Fast image dehazing using guided filter," in *Proc. IEEE 16th Int. Conf. Commun. Technol. (ICCT)*, Oct. 2015, pp. 182–185.
[45] J. Xu, W. Zhao, P. Liu, and X. Tang, "Removing rain and snow in a single image using guided filter," in *Proc. IEEE Int. Conf. Comput. Sci. Autom. Eng. (CSAE)*, vol. 2, no. 2, May 2012, pp. 304–307.
[46] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2864–2875, Jul. 2013.
[47] T. Jinno and M. Okuda, "Multiple exposure fusion for high dynamic range image acquisition," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 358–365, Jan. 2012.
[48] W. Zhang and W.-K. Cham, "Gradient-directed multiexposure composition," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2318–2323, Apr. 2012.
[49] T.-H. Oh, J.-Y. Lee, Y.-W. Tai, and I. S. Kweon, "Robust high dynamic range imaging by rank minimization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 6, pp. 1219–1232, Jun. 2015.
[50] A. Srikantha and D. Sidibé, "Ghost detection and removal for high dynamic range images: Recent advances," *Signal Process., Image Commun.*, vol. 27, no. 6, pp. 650–662, Jul. 2012.
[51] G. H. Turner and D. B. Twieg, "Study of temporal stationarity and spatial consistency of fMRI noise using independent component analysis," *IEEE Trans. Med. Imag.*, vol. 24, no. 6, pp. 712–718, Jun. 2005.
[52] D. J. McFarland, L. M. McCane, S. V. David, and J. R. Wolpaw, "Spatial filter selection for EEG-based communication," *Electroencephalogr. Clin. Neurophysiol.*, vol. 103, no. 3, pp. 386–394, Sep. 1997.
[53] R.-J. Li, S.-Y. Yu, and X.-W. Wang, "Unsupervised spatio-temporal segmentation for extracting moving objects in video sequences," *J. Shanghai Jiaotong Univ. (Sci.)*, vol. 14, no. 2, pp. 154–161, Apr. 2009.
[54] K. Ma, K. Zeng, and Z. Wang, "Perceptual quality assessment for multi-exposure image fusion," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3345–3356, Nov. 2015.
[55] C. Yang, J.-Q. Zhang, X.-R. Wang, and X. Liu, "A novel similarity based quality metric for image fusion," *Inf. Fusion*, vol. 9, no. 2, pp. 156–160, Apr. 2008.
[56] E. G. Turitsyna and S. Webb, "Simple design of FBG-based VSB filters for ultra-dense WDM transmission," *Electron. Lett.*, vol. 41, no. 2, pp. 40–41, 2005.
[57] Y. Chen and R. S. Blum, "A new automated quality assessment algorithm for image fusion," *Image Vis. Comput.*, vol. 27, no. 10, pp. 1421–1432, Sep. 2009.
[58] C. S. Xydeas and V. S. Petrovic, "Objective pixel-level image fusion performance measure," *Proc. SPIE*, vol. 4051, pp. 89–98, Apr. 2000.
[59] L. Krasula, P. Le Callet, K. Fliegel, and M. Klima, "Quality assessment of sharpened images: Challenges, methodology, and objective metrics," *IEEE Trans. Image Process.*, vol. 26, no. 3, pp. 1496–1508, Mar. 2017.
[60] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganiere, and W. Wu, "Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 1, pp. 94–109, Jan. 2012.
[61] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electron. Lett.*, vol. 38, no. 7, pp. 313–315, Mar. 2002.

MUHAMMAD SOHAIB ROOMI received the M.Sc. degree in information technology from the University of Sargodha, Pakistan, in 2015. He is currently pursuing the M.S. degree in computer science with the Balochistan University of Information Technology, Engineering and Management Sciences, Quetta, Pakistan. His current research areas include high dynamic range images, image fusion, and image de-ghosting.

MUHAMMAD IMRAN received the Ph.D. degree in electrical engineering from Florida State University, USA, in 2016. He is currently an Assistant Professor with the Department of Electrical Engineering, Balochistan University of Information Technology, Engineering and Management Sciences, Quetta, Pakistan. His research interest includes image fusion, multi-resolution decomposition, high dynamic range images, and image de-ghosting.

SYED ATTIQUE SHAH received the Ph.D. degree from the Institute of Informatics, Istanbul Technical University, Istanbul, Turkey. During his Ph.D., he was a Visiting Scholar with National Chiao Tung University, Taiwan, The University of Tokyo, Japan, and the Tallinn University of Technology, Estonia, where he completed the major content of his thesis. He currently works as an Assistant Professor and the Chairperson with the Department of Computer Science, Balochistan University of Information Technology, Engineering and Management Sciences, Quetta, Pakistan. His research interests include big data analytics, cloud computing, information management, and the Internet of Things.

AHMAD ALMOGREN (Senior Member, IEEE) received the Ph.D. degree in computer science from Southern Methodist University, Dallas, TX, USA, in 2002. He was the Vice Dean of the Development and Quality, College of Computer and Information Sciences (CCIS), King Saud University (KSU), Riyadh, Saudi Arabia, where he is currently a Professor and the Director of the Cyber Security Chair with the Department of Computer Science. He was the Dean of the College of Computer and Information Sciences and the Head of the Academic Accreditation Council, Al-Yamamah University. His research interests include mobile-pervasive computing and cybersecurity. He served as the General Chair for the IEEE Smart World Symposium and a Technical Program Committee Member in numerous international conferences/workshops, such as the IEEE CCNC, ACM, BodyNets, and the IEEE HPCC.

IHSAN ALI received the M.S. degree in computer system engineering from Ghulam Ishaq Khan Institute of Engineering Sciences and Technology in 2008. He is currently pursuing the Ph.D. degree with the Faculty of Computer Science and Information Technology, University of Malaya. He has been actively involved in research and teaching activities for the last ten years in the different country including Saudi Arabia, USA, Pakistan, and Malaysia. He is currently an Active Research Associate with the Centre for Mobile Cloud Computing Research, Faculty of Computer Science and Information Technology, University Malay, Kuala Lumpur, Malaysia. He has authored or coauthored more than 40 high impact research journal articles, including a highly reputable IEEE COMMUNICATIONS SURVEYS & TUTORIALS and the *IEEE Communication Magazine*. His research interests include wireless sensor networks (WSNs), robotics in WSNs, sensor cloud, fog computing, the IoT, ML/DL in WSN. He has served as a Technical Program Committee Member for several well-known conferences, including the IWCMC 2017–2018, AINIS 2017, Future 5V 2017, ICACCI-2018, INAIT2019, DiCES-N19, CCNC2020, ICCAIS2020, and CSNT2020, and also an organizer of the special session on fog computing in Future 5V 2017. He is also an Active Reviewer of *Computers & Electrical Engineering*, the *KSII Transactions on Internet and Information Systems*, *Mobile Networks and Applications*, the *International Journal of Distributed Sensor Networks*, the *Journal of Advanced Transportation*, the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, COMPUTER NETWORKS, IEEE ACCESS, *Wireless Communications and Mobile Computing*, and the *IEEE Communication Magazine*. Detail profile can be seen at https://orcid.org/0000-0002-9549-2540.

MANSOUR ZUAIR received the B.S. degree in computer engineering from King Saud University and the M.S. and Ph.D. degrees in computer engineering from Syracuse University. He has served as the Chairman with CEN from 2003 to 2006 and the Vice Dean from 2009 to 2015, where he has been the Dean since 2016. He is currently an Associate Professor with the Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. His research interests include computer architecture, computer networks, and signal processing.