IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 7, 2020, accepted August 17, 2020, date of publication August 20, 2020, date of current version September 9, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3018264

An Enhanced Image Fusion Algorithm by Combined Histogram Equalization and Fast Gray Level Grouping Using Multi-Scale Decomposition and Gray-PCA

JAMEEL AHMED BHUTTO[1], TIAN LIANFANG[1,2,3], (Senior Member, IEEE), QILIANG DU[1,2,3], TOUFIQUE AHMED SOOMRO[4], (Member, IEEE), YU LUBIN[1], AND MUHAMMAD FAIZAN TAHIR[5]

[1]School of Automation, South China University of Technology, Guangzhou 510640, China
[2]Research Institute of Modern Industrial Innovation, South China University of Technology, Guangzhou 510640, China
[3]Key Laboratory of Autonomous Systems and Network Control of Ministry of Education, Guangzhou 510640, China
[4]Department of Electronic Engineering, Quaid-e-Awam University of Engineering, Science and Technology, Nawabshah 67480, Pakistan
[5]School of Electric Power, South China University of Technology, Guangzhou 510640, China

Corresponding authors: Tian Lianfang (tialiangfang@scut.edu.cn) and Qiliang Du (qldu@scut.edu.cn)

This work was supported in part by the Department of Natural Resources of Guangdong Province-Offshore Wind Power Project 2020, in part by the Key-Area Research and Development Program of Guangdong Province under Grant 2019B020214001, in part by the Key-Area Research and Development Program of Guangdong Province under Grant 2018B010109001, in part by the Guangdong Industrial Technology Major Research Plan under Grant 2019-01-01-12-1006-0001, and in part by the Fundamental Research Funds for the Central Universities under Grant 2018KZ05.

ABSTRACT Image enhancement is a challenging task in image analysis particularly, it is more challenging in performing image fusion. Image fusion is the process of combining multiple images to produce quality output without any variation in contrast, blurring, and noise. Many image fusion algorithms have been implemented, but their final fused images suffer from variations in background contrast, uneven illumination, blurring, and the presence of noise. To overcome the aforementioned issues, this paper proposed a new image fusion method, which improves image contrast and also gives appropriate details of the image. Our method is based on a set of conventional techniques such as amalgamated histogram equalization and fast gray-scale grouping to handle the problems mentioned, and we improve overall fusion strategies by proposing a novel principal component analysis technique to convert RGB types images to high gray-scale contrast image as the final output image. We have carried out many experiments on different common databases used by various researchers. Our proposed method gives good subjective and objective performances compared to other statuses. Our proposed method can be used in different monitoring applications.

INDEX TERMS Histogram equalization and fast gray-level grouping (HEFGLG), image fusion, multi-focus image, infrared image, visible image, average pixel intensity, cross correlation.

I. INTRODUCTION

The image fusion is the way of merging multiple input images to produce one output image which extracts high quality and more informative for the perception of human vision, robot and other processing tasks as compared to any of the input images [1]–[3]. In recent years, image fusion techniques have become a promising research area and have received a lot of interest in many applications such as computer vision, face detection and recognition, medical diagnose, surveillance, and so on [1], [4]–[6].

The image of different types such as computed tomography (CT), magnetic resonance imaging (MRI), visible, infrared and the images which are taken by the same camera with different focal lengths are suitable for fusion of images [7]. However, one model cannot capture enough information due to limitations of the system because it is hard to capture focus for all objects due to the limited focal length of camera. Therefore, combining images of different focal lengths for the same scenery by varying the focal length is known as

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

multi-focus image fusion [8]–[11]. Similarly, CT and MRI images are used to diagnose many medical conditions such as strokes, tumors. However, CT images provide only information about bones, whereas MRI images give information about soft tissues, so we cannot get sufficient information from a single modal medical image. Therefore, image fusion is used in this case to merge both CT and MRI images to get information of soft tissues and clear bones, which is known as medical image fusion [12]–[15]. Similarly, the visible images (VI) capture the image details when there is enough illumination. However, the contrast is poor when there is insufficient brightness. On the other hand, infrared images (IR) cannot capture the real details of scene because it can only reflect the gray-scale of the scene [16], [17]. Consequently, an image fusion technique is used in this perspective to get complementary information from both images which can be used in surveillance applications [1]–[3].

The image fusion is performed by three distinct processing levels named as pixel-level, feature-level and decision-level [5], [18]–[21]. Pixel-level based fusion is used by many researchers in various applications. It merges the pixels of input images directly to acquire the final output image [1], [3]. Some researchers have also taken attention on feature-level based image fusion that deals with high-level processing tasks. It extracts the features in the image and then fuses the features using some advanced fusion schemes such as region-based [1], [2]. The purpose of feature-level fusion is to acquire desired features from input images instead of all features [1], [2]. The decision level is the highest processing level of three levels mentioned above. It extracts all information from images and then according to specific criteria, the decisions are taken to fuse the extracted information [1]. This type of processing is widely used in biometrics, fingerprint verification and face recognition.

The main key goal of any image fusion algorithm is to increase contrast of source images so that it can preserve most of the useful information without producing artifacts and apply proper fusion strategy that should be robust to improper conditions. Though many image fusion algorithms have been designed till now but the fused image suffers from variations in background contrast, uneven illumination, blurring effect and the presence of noise. To overcome the shortcomings of the aforementioned issues, this paper presents an enhanced image fusion algorithm that addresses these issues and achieves better results than existing state-of-art image fusion techniques.

The main contributions of this work are as follows:

i) An amalgamated histogram equalization and fast gray-level grouping (HEFGLG) technique is proposed in this work that automatically enhances the contrast of image. This method is computationally efficient and it also produces better results.

ii) Non-subsampled contourlet transform (NSCT) is applied to images, which are obtained from HEFGLG. This decomposition method can efficiently capture the true geometrical structure. Moreover, it has shift-invariance property with fast computation.

iii) Local energy-based fusion strategy by choosing either averaging or selection mode is applied for low-frequency images that preserve the energy information. Mean-gradient fusion strategy is applied for high-frequency images to maintain the texture, edges, boundaries and smooth contours.

iv) Principal component analysis (PCA) based multi-channel color to single-channel gray conversion technique is implemented for dimension reduction and robust operation. This technique effectively preserves the discriminability between color and textures in the resultant image.

The rest of the paper is organized as follows. Section 2 summarizes the literature review of recent image fusion algorithms. Section 3 elaborates the proposed work. Experimental results and evaluation of algorithm are presented in section 4, and the conclusion is stated in section 5.

II. THE RELATED WORK

Image fusion is the most popular research area and it has received remarkable progress in last few decades. We present literature on recent studies for state-of-art image fusion techniques.

The PCA is applied in [22], which reduces the dimensions of data and produces energy information in final fused output image [18]. Nevertheless, this method suffers from spectral degradation, and due to that, it cannot capture the smooth edges, textures and contours in the image. Discrete wavelet transform (DWT) was proposed by Yon *et al.* [23], which preserves high-frequency information with fast computation but it has lack of shift-invariance property that introduce artifacts and noise. Moreover, it has limited directional information and cannot capture the essential features like contours and edges. Mane *et al.* [13] introduced hybrid DWT and PCA so that it can achieve spatial as well as spectral information. The input images are decomposed by DWT and then PCA is applied to decomposed images. Though this method produces better fused results than individual DWT and PCA techniques but the output image still has limited directional information. Moreover, this combined technique is shift variant that introduces artifacts in the fused image and it also cannot capture the smooth contours and edges. To address the aforementioned issues, contourlet transform is used by two different authors in [19], [24]. It can capture smooth edges and contours, and gives multi-directional information. However, this method uses downsampling that makes it shift variant and it leads to Gibb's effect; thus, it affects the result of output image. The NSCT is discussed in [25], which is fully shift-invariant and has flexible frequency selectivity with fast implementation that addresses the above issues. However, the poor illumination and blurring effect degrades the fused image so it cannot preserve the detailed information from input images. The edge-preserving filter based image fusion is proposed by Tian *et al.* [4] using combined median-average based discrete stationary wavelet

transform (DSWT) with PCA. At first, the median-average filter eliminates the noise from each pixel. Then, DSWT is applied to decompose the images and finally, PCA is used for data reduction. This method achieves better results but the fused image still suffers due to limited directional information and it cannot capture smooth edges and contours. The hybrid non-subsampled shearlet transform (NSST) based spatial frequency (SF) and pulse coupled neural network (PCNN) is implemented in [26]. The NSST is used to decompose both images and SF-PCNN is applied to sub-band coefficients for source images. This algorithm addresses the issues of shift variance and it has better frequency selectivity. However, the decomposed low-frequency image is affected by poor contrast and sharpness, so the image loses some part of energy information, which affects the overall performance of the output image [27]. An improved morphology-hat transform (MT) based image fusion algorithm using contourlet transform (CT) and PCA has been designed by He Li et al in [28]. Though, this algorithm increases the brightness of fused image, preserves the energy information and captures smooth contours. However, contourlet transform is shift variant because it uses down-sampling and thus Gibbs effect is introduced. In [29], the author has implemented fully convolutional network for the multi-focus image fusion method. In this work, the pooling layers are adjusted to change in convolutional layers; therefore, all layers are convolutional layers. The better visual effect is obtained by this method. However, it is very time consuming and it is also very hard to design the structure of network and train data sets. Shuaiqi et al in [10] have designed image fusion algorithm by amalgamating NSST and residual network (ResNet). It first decomposes the input images by NSST and then ResNet is applied for low frequency images, and enhanced gradient sum of Laplacian energy is performed on high frequency images. The fused image obtained by this method produces more clear information and better details. However, this algorithm is time consuming. The author in [30] apply the Latent low-rank representation (LatLRR) to fuse the infrared and visible images. However, the artifacts in the edge areas are introduced due to lack of spatial consistency in this method.

III. THE PROPOSED WORK

The existing image fusion algorithms have their pros and cons, and their pros should be amalgamated to enhance the quality of the fused image. There is need to design an image fusion algorithm that can automatically adjust the contrast of images so that it can preserve energy information, smooth edges, contours and sufficient complementary information without introducing artifacts and noise. Besides, apply the improved fusion strategy that should be robust to improper conditions while achieving high quality fused image. Moreover, the fused image can effectively preserve discriminability between color and textures in the resultant image by using some data reduction technique with the robust operation. This paper presents an enhanced image fusion algorithm that meets the aforementioned requirements and achieves better results than existing image fusion methods.

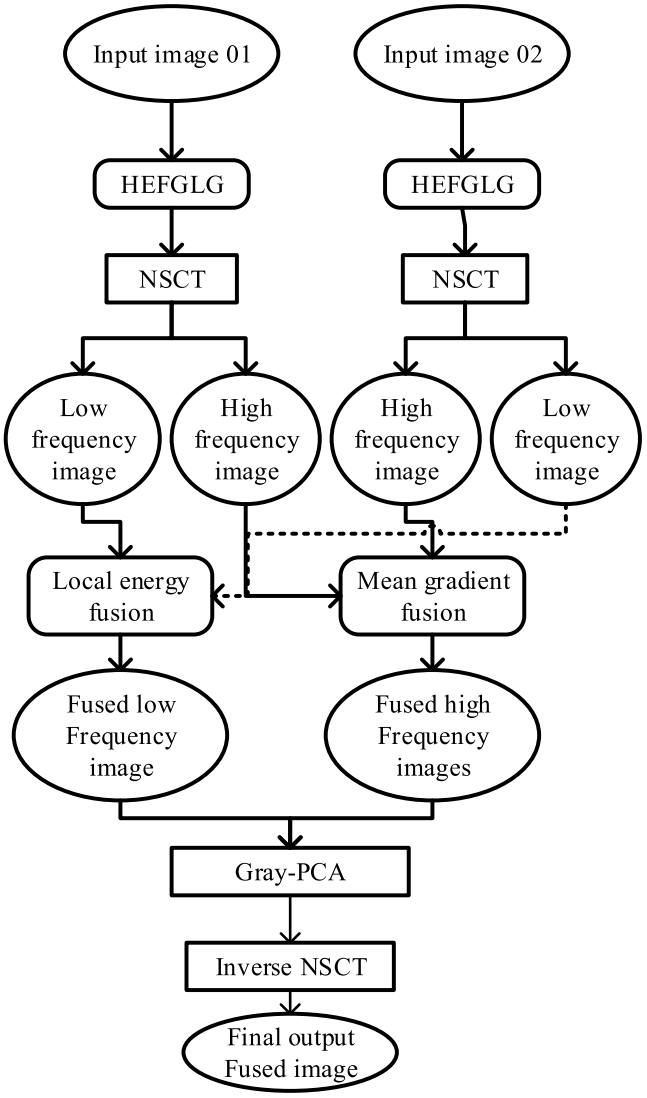

FIGURE 1. Block diagram of the proposed method.

The proposed method includes several sequential stages as depicted in Figure 1. The main goal of each stage is to enhance the quality of an image without introducing artifacts. Each stage is explained in details in subsequent sub-sections:

A. COMBINED HEFGLG

It is challenging to enhance the low contrast images either natural, multi-focus, infrared and visible or medical images whose intensity of image in gray domain is very high at one position and very small in other parts of image. To overcome the above challenges, HEFGLG is used in this paper. It divides the histogram of low varying contrast into two histograms in accordance with the position of the highest intensity histogram component. Histogram equalization (HE) is used to increase contrast by equalizing left segment of

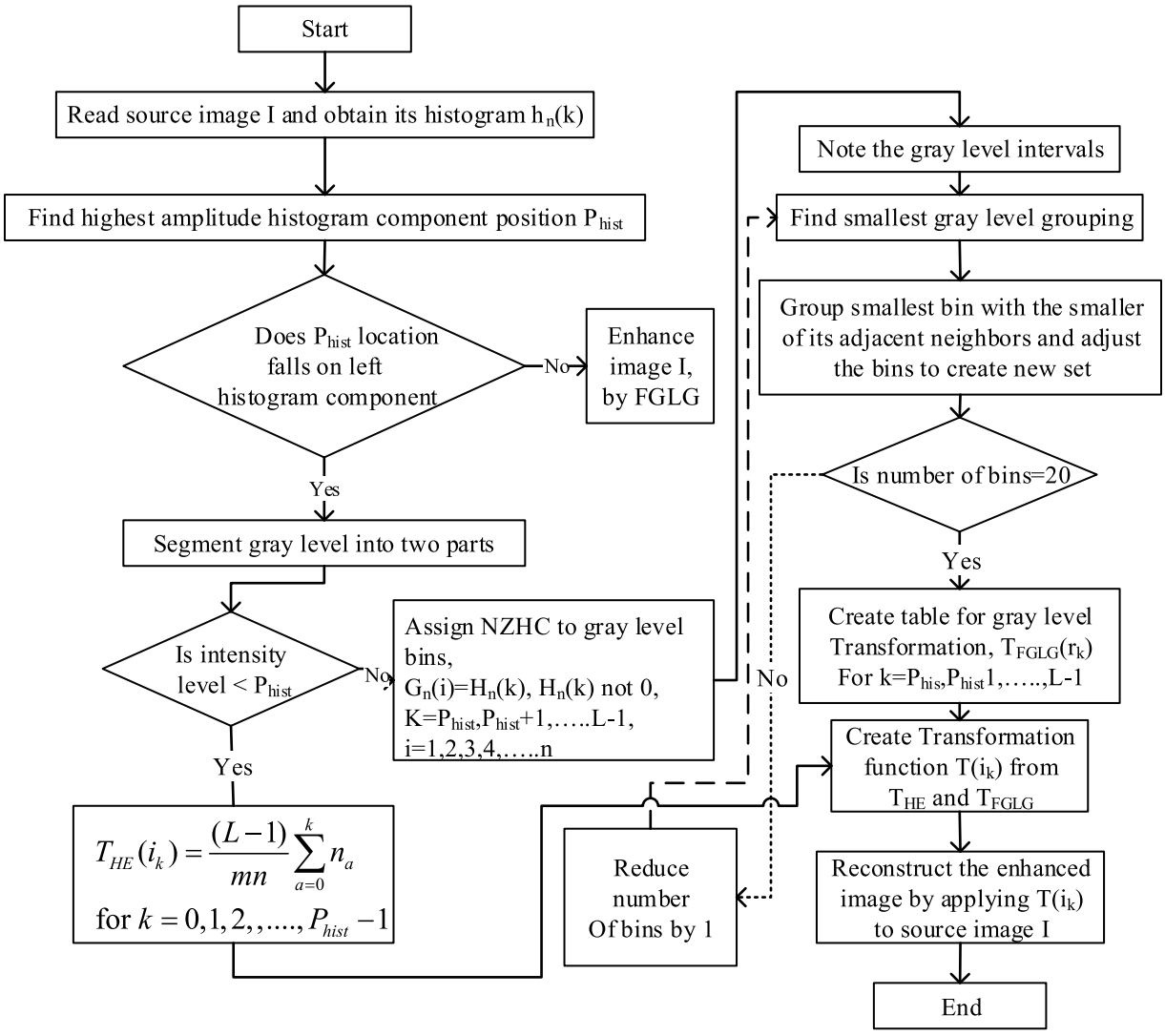

FIGURE 2. Flowchart of histogram equalization and fast gray level grouping.

histogram component and fast gray level grouping (FGLG) is applied to histogram component for right side [31].

As gray level grouping (GLG) [32] technique first group the parts of low contrast image into the desired number of groupings based on predefined criteria. Furthermore, it redistributes these histogram components groups uniformly so that each group hold same size gray-scale segment as in other groups. Finally, all of the previously grouped gray levels are ungrouped. This whole process is time-consuming, increases the complexity and also produce washed-out effect that degrades the quality of image. This paper employs FGLG that uses a default value for all gray level bins [33] that makes it fully automated as it does not require the construction of the transformation function and calculation of the average distance between two pixels for each set of gray-levels. We set 20 as a default value for gray level bins that can be seen on right side of Figure 2, which reduces the time and the number of iterations. This feature is not only computationally efficient as it requires less time and number of iterations but also produces better results. Figure 2 shows the schematic diagram of HEFGLG.

The histogram of an input image $h(i_k)$ is a discrete function $h(i_k) = n_k$ with the intensity level in the range of $[0, L-1]$, where i_k is the kth amplitude level and n_k is the total pixel numbers for input image. The histogram is normalized by dividing its all components using total pixel numbers denoted by product of $m \times n$ in an image, where m and n indicate rows and columns of that image. Therefore normalized histogram is given by $P(i_k) = n_k/mn, k = 0, 1, 2, 3, \ldots\ldots, L-1$ and $P(i_k)$ is the probability of occurrence for i_k in the source image. After the histogram is calculated for the source images with intensity level in range of $[0, L-1]$, the procedure of HEFGLG is as follows:

(1) Find the highest amplitude histogram component position P_{hist} on gray-scale for the source image. If the P_{hist} lies on left side but not in the first component of NZHC, then the

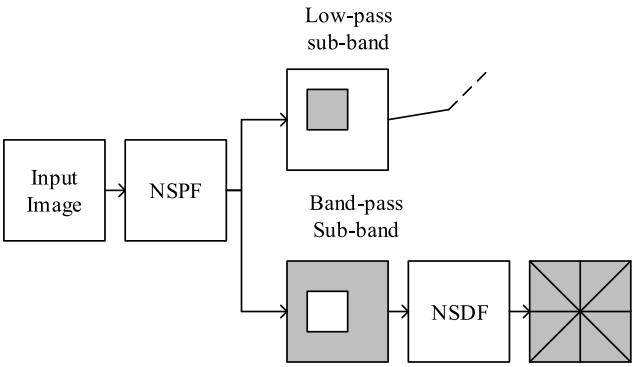

FIGURE 3. Block diagram of NSCT decomposition [25].

original histogram is segmented into two histograms, the first start from (0 to $P_{hist}-1$) and other from (P_{hist} to $L-1$). If the P_{hist} lies either inside the right part or in first component of NZHC, then simply improve the low contrast of input image using FGLG.

(2) Apply HE in the first histogram component (0 to $P_{hist}-1$), and use FGLG for other histogram componen (P_{hist} to $L-1$).

(3) The transformation function HE is computed as

$$T_{HE}(i_k) = (L-1)\sum_{a=0}^{k} P_i(i_a) = \frac{(L-1)}{mn}\sum_{a=0}^{k} n_a \qquad (1)$$

Here $k = 0, 1, 2, 3, \ldots, P_{hist}-1$, mn is number of rows and columns, and the transformation function for FGLG is computed as:

$$T_{FGLG}(i_k), \quad \text{for } k = P_{hist}, P_{hist}+1, P_{hist}+2, P_{hist}+3, \ldots\ldots, L-1.$$

(4) Therefore, the piecewise transformation function is calculated as:

$$T(i_k) = T_{HE} + T_{FGLG}(i_k) \qquad (2)$$

The piecewise transformation function is applied for reconstruction to get an enhanced and rich contrast image.

B. NSCT

The NSCT is multiscale and multi-directional image decomposition method designed by amalgamating the non-subsampled pyramid filter banks (NSPFB) and non-subsampled directional filter banks (NSDFB) [34]. This decomposition method can efficiently restore the true geometrical structure of an image such as edges, contours, and it is fully shift-invariant.

The NSPFB decomposes the image by using two-channel non-subsampled 2D filter banks by avoiding down-sampling or up-sampling. Further, NSDFB is used to split band-pass sub-bands into different directions, as depicted in Figure 3.

The perfect reconstruction for NSPFB is obtained by:

$$H_0(z)G_0(z) + H_1(z)G_1(z) = 1 \qquad (3)$$

where $H_1(z) = 1 - H_0(z)$; $G_0(z)$ and $G_1(z)$ are low-pass and band-pass filters.

The non-subsampled pyramids are designed using iterative non-subsampled filter banks to obtain multiscale decomposition. Then, all filters are up-sampled by 2 for the next level and due to that, they satisfy the perfect reconstruction criteria. The equivalent filter for $k-th$ level cascaded non-subsampled pyramids is expressed as:

$$H_n^{equ}(z) = \begin{cases} H_1(z^{2n-1})\prod_{j=0}^{n-2} H_0(z^j) & 1 \le n \le 2^k \\ \prod_{j=0}^{n-1} H_0(z^{2j}) & n = 2^k \end{cases} \qquad (4)$$

Here z^j stands for $[z_1^j, z_2^j]$. After applying decomposition by NSPFB, the NSDFB is applied that is a shift-invariant [35]. It is also two-channel non-subsampled filter bank iterative method which achieves better directional decomposition. This same process will be used for desired higher level decomposition.

C. FUSION STRATEGY FOR LOW-FREQUENCY AND HIGH-FREQUENCY IMAGES

The performance of the final fused image is directly affected by the fusion strategy and the way it is applied to images. At present, a weighted average fusion strategy is used but it cannot capture useful features such as energy information, textures, edges and contours in the fused image.

In this paper, local energy based fusion strategy and mean-gradient based fusion strategy are applied for low-frequency and high-frequency images. The accuracy in the fusion of low-pass images can be further improved with a silence feature which determines whether selection or averaging mode is applied. Meanwhile, the mean-gradient based fusion strategy is employed to high-pass images that can efficiently capture the smooth edges, contours, and boundaries.

1) FUSION STRATEGY FOR LOW-FREQUENCY IMAGES

The low-frequency images have mainly the energy information and it is vital to preserve the contrast details of input images. The simplest way is to apply averaging fusion technique but it cannot capture high-quality images. Therefore, local energy based fusion strategy is applied for NSCT, and then two distinct combination modes; averaging and selection modes are used for coefficients to enhance the fusion accuracy.

First, the local energy $E_l(x, y)$ is computed by centering the current coefficients in the coarse sub-band C_J that is given as:

$$E_l(x, y) = \sum_m \sum_n C_J(x+m, y+n)^2 W_L(m, n) \qquad (5)$$

Here (x, y) represent the current non-subsampled contourlet coefficients and W_L is a template of size 3×3.

The silence factor S_J is computed to decide whether averaging or selection mode is used in the fusion process. The

silence factor S_J is given by:

$$S_J^{AB}(x,y) = \frac{2\sum_m \sum_n C_J^A(x+m, y+n) C_J^B(x+m, y+n)}{E_l^A(x,y) + E_L^B(x,y)} \quad (6)$$

Here $C_J^X(x,y), X = A, B$ represent the low pass non-subsampled contourlet coefficients for input image A or B, respectively.

The silence factor S_J reflects the similarity between two input images for low-pass sub-bands. The S_J value is compared with threshold level T. If $S_J^{AB} > T$ the input coefficients in c_J^A and c_J^B are very similar and averaging mode is applied, and the information is obtained from both input images. Its equation is computed by:

$$c_J^F(x,y) = \beta_A \cdot c_J^A(x,y) + \beta_B \cdot c_J^B(x,y) \quad (7)$$

c_J^F denotes the final fused coefficients at the position (x, y), β_A and β_B are weights.

$$\beta_A = \begin{cases} \beta_{\min} & for\ E_l^A(x,y) < E_l^B(x,y) \\ \beta_{\max} & for\ E_l^A(x,y) \geq E_l^B(x,y) \end{cases} \quad (8)$$

$$\beta_B = 1 - \beta_A \quad (9)$$

Here $\beta_{\min} \in (0, 1)$, $\beta_{\min} + \beta_{\max} = 1$

On the other hand, if $S_J^{AB} \leq T$ it corresponds to the dissimilarity between the input coefficients c_J^A and c_J^B, the selection mode is applied in this case. The coefficients with larger energy are selected, while coefficients with small energy are discarded. The selection mode is computed by:

$$\beta_J^F = \begin{cases} \beta_J^A(x,y) & for\ E_l^A(x,y) \geq E_l^B(x,y) \\ \beta_J^B(x,y) & for\ E_l^A(x,y) < E_l^B(x,y) \end{cases} \quad (10)$$

Here $E_l^A(x,y)\ and\ E_l^B(x,y)$ represents the local energy of image A *and* B.

2) FUSION STRATEGY FOR HIGH-FREQUENCY IMAGES

The high-frequency images contain the textures, edges, contours and object boundaries of original image. This paper incorporates the mean-gradient fusion scheme.

The mean gradient (G) of region (R) is computed first by:

$$G = \frac{1}{(M-N)} \sum_{x=1}^{M-1} \sum_{y=1}^{N-1} \sqrt{\frac{\Delta I_x^2 + \Delta I_y^3}{2}} \quad (11)$$

The size of R is $M \times N\,(M, N \text{ is odd}, N \geq 3, M \geq 3)$, ΔI_x and ΔI_y are the first-order difference of $f(x,y)$ in X and Y directions.

$$\Delta I_x = f(x,y) - f(x-1,y) \quad (12)$$

$$\Delta I_y = f(x,y) - f(x,y-1) \quad (13)$$

After that, fusion coefficient (β) of two images is obtained by mean gradient of both input images, which is given by:

$$\beta_A = \frac{G_A}{G_A + G_B}, \quad \beta_B = \frac{G_B}{G_A + G_B} \quad (14)$$

where β_A and β_B are the mean gradient of both images.

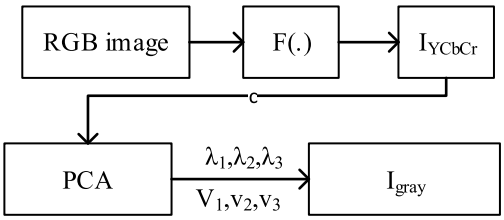

FIGURE 4. Schematic diagram of PCA-based gray conversion scheme.

Finally, the fused coefficients for high-frequency images are computed by:

$$F(x,y) = \beta_A f(x,y) + \beta_B f(x,y) \quad (15)$$

D. PRINCIPAL COMPONENT ANALYSIS (PCA) BASED GRAY CONVERSION METHOD

The next essential part is to apply a gray-PCA technique that reduces the amount of data and produce an enhanced single-channel gray output image with fast computation and robust operation [36]. Figure 4 shows the schematic diagram of the gray-PCA based color-to-gray conversion process.

The image that is processed must be a luminance-chrominance representation but if the image is in RGB color space, then we need to convert it into luminance-chrominance representation. The first step is formation of the vectorized color image ($I_{rgb} \in R^{3\times n}$) by stacking image representation into RGB color channels side-by-side. Subsequently, zero-mean (YC_bC_r) processed image ($I_{YC_bC_r} \in R^3$, Y and C_bC_r indicate the luminance and chrominance) is obtained by separating the luminance and chrominance using a transfer function $f(.)$, shown in paper [37].

After that, Eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \in R^1$ and their normalized eigenvectors $v_1 \geq v_2 \geq v_3 \in R^3$ are computed by applying principal component analysis. The resultant single-channel gray-scale image I_{gray} is computed by weighted linear combination of three projections, and the weights are obtained by percentage of their eigenvalues. As a result, the first subspace projection results in multicolor-to-gray mapping because of its highest eigenvalue. In contrast, the second and third subspace projection preserves the details for colored images in the resulting single-channel gray-scale image.

IV. EXPERIMENTAL RESULTS, DATASETS AND EVALUATION OF ALGORITHM

We have taken ten pairs of input images and two pairs of input images are processed separately by each stage to clearly show the influence of each stage in the proposed method. The two pairs of input images are shown in Figure 5. The HEFGL is adopted to enhance the contrast of images. It can be clearly seen in Figure 6 that HEFGL method enhances the contrast of input images and images are more vivid with sufficient information. The enhanced images are then processed by NSCT that uses four multiscale decomposition levels. The NSCT has property of shift-invariance with flexible frequency selectivity that results in high quality image

FIGURE 5. Two pairs of input images.

FIGURE 6. The images obtained by HEFGLG.

without introducing artifacts. The images generated by NSCT are depicted in Figure 7. Finally, the gray-PCA method is applied to images that are obtained from NSCT. The gray-PCA method not only reduces the dimensions of data with robust operation but this method also effectively preserves the discriminability between color and textures in the resultant image and produces high quality output image. The final fused image is depicted in Figure 8. Therefore, it can be analyzed that images generated at each stage produce better quality image, which shows the positive influence of each stage in proposed method.

A. DATASETS

In this paper, we have used ten pairs of datasets for RGB images to compare proposed method with state-of-the-art

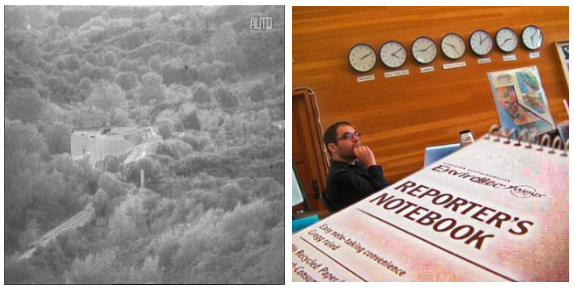

FIGURE 7. The fused images of NSCT.

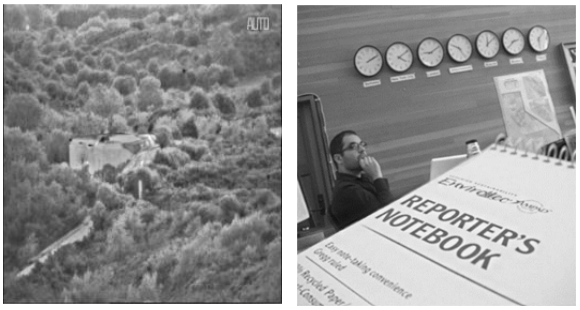

FIGURE 8. The final fused images after applying gray-PCA.

image fusion techniques. Many authors have used these datasets for image fusion research field. For instance, the medical images which contain registered CT and MRI are publically distributed by the Harvard Medical school at http://www.med.harvard.edu/AANLIB/home.html, and McConnel Brain Imaging Centre of the Montreal Neurological Institute has distributed datasets at http://www.mouldy.bic.mni.mcgill.ca/brainweb. These datasets have been applied in several image fusion areas [38]–[41]. The datasets for infrared and visible imaging are taken from http://www.metapix.de/indexp.htm, http://ece.lehih.edu/SPCRL/IF/image_fusion.html, and http://web.media.mit.edu/∼ raskar/NPAR04/. These datasets have been used in several image fusion research areas [7], [39], [42]–[44]. The multi-focus datasets are available at https://dsp.etfbl.net/mif/, which is shared by Savic and it has been used in different papers for image fusion [45], [46].

All source images are RBG but they are categorized into two scenarios: RGB images visualized as colored image and RGB images visualized as a gray image. The results of all final fused images are converted from three-channel RBG image to single-channel gray image.

1) SCENARIO ONE: RGB IMAGES VISUALIZED AS GRAY IMAGES

In this scenario, we use the RGB source images that are visualized as gray images. Figure 9 to 14 shows the RGB source images that visualize as gray images and the results of final fused images.

2) SCENARIO TWO: RGB IMAGES VISUALIZED AS COLORED IMAGES

In this scenario, we use the RGB source images that are visualized as colored images. Figure 15 to 18 shows the RGB

source images that visualize as colored images and the results of final fused images.

B. EVALUATION OF ALGORITHM

Evaluating the performance of the proposed algorithm with existing techniques is a challenging task [47]. The performance evaluation is categorized in subjective and objective evaluation. We have compared the experimental results with PCA [22], DWT [23], median-average based DSWT-PCA [4], non-subsampled shearlet transform based spatial frequency and pulse code neural network (NSST-SF-PCNN) [26] and morphology-hat transform based contourlet transform with principal component analysis (MT-CT-PCA) [28] image fusion techniques. The median-average based DSWT-PCA, NSST-SF-PCNN and MT-CT-PCA are new hybrid image fusion algorithms, which perform better than other techniques but the proposed method achieves the best performance among the aforementioned schemes.

1) SUBJECTIVE EVALUATION

The subjective evaluation is used to evaluate the quality of the final fused image according to the perception of human vision. It is the most widely used and popular way to visualize the difference between images by the human eye. Figure 9 to 18 shows the subjective evaluation of proposed method and five aforementioned state-of-art image fusion techniques.

It can be seen in Figure 9, the fusion effect of the NSST-SF-PCNN method, median-average and DWT is almost similar. Their contrast is high but the details of these fused images are not precise; some useful information about soft tissues is missing. Compared to NSST-SF-PCCN, median-average and DWT fusion methods; the PCA produces better information about soft tissues. But the fused image has washed out effect due to which the combined image has artifacts that degrade the overall quality of fused image. In comparison to these fusion schemes, the MT-CT-PCA method has better contrast and improvement in the information about soft tissues. However, the overall effect is still not satisfactory. It can be seen that the proposed method adjusts the contrast by using HEFGLG and we can get precise information about soft tissues highlighted in red boxes of image. The overall effect of the proposed method is much better than aforementioned methods.

Figure 10 shows the fused results for multi-focus clock images. It can be clearly seen that proposed method achieves much better results for fused images than existing techniques. The brightness of the clock in proposed method is consistent with the clock in the source images. Moreover, it is obvious in red highlighted boxes that the contours are precise; the curve and edges in left clock are smooth with bright contrast. Though the MT-CT-PCA fusion scheme also has better contrast but the curve and boundaries are not stable as in the proposed method, which is highlighted in Figure 10 with red boxes. The performance of NSST-SF-PCNN fusion, median-

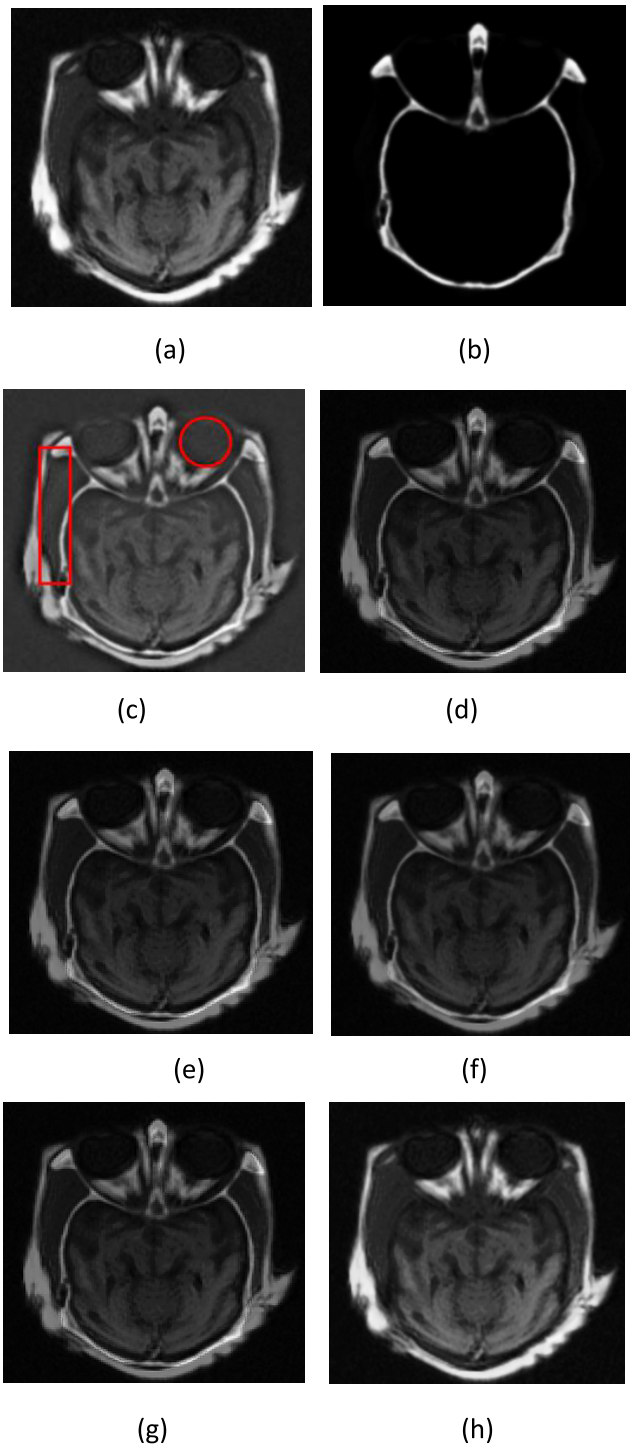

FIGURE 9. Input images and fused images. (a) CT image; (b) MRI image; (c) proposed method; (d) MT-CT-PCA; (e) NSST-SF-PCNN; (f) median-average-DSWT-PCA; (g) DWT; (h) PCA.

average and DWT almost resembles; the contrast is not bright, the image is blurred and overall results are not satisfactory.

The fusion effect of MT-CT-PCA, NSST-SF-PCNN fusion and DWT in Figure 11 is not satisfactory. The contrast is also very poor and the objects in the image have no clarity. The median-average method produces better results than existing

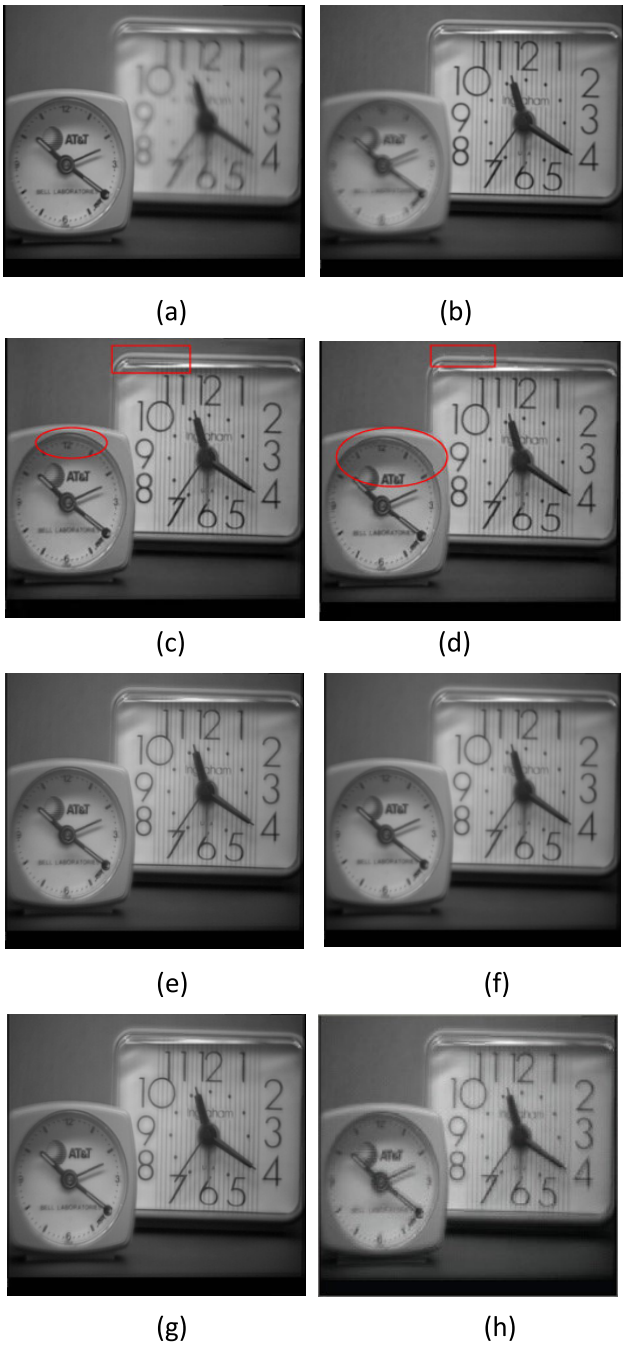

FIGURE 10. Multi-focus clock input images and fused images. (a) & (b) are input images; (c) proposed method; (d) MT-CT-PCA; (e) NSST-SF-PCNN; (f) Median- average-DSWT-PCA; (g) DWT; (h) PCA.

techniques and three persons are visible. However, the gun and car wheel are not visible in the median-average method and the contrast of fused image is also not consistent with input images. It can be analyzed in Figure 11 that proposed method produces much better results. Three persons, car wheel and the gun are more vivid in the proposed method. Furthermore, the brightness of the three persons and the gun is consistent with the source images.

The fusion effect of Infrared and visible image for desert-car is depicted in Figure 12. The fusion effect of

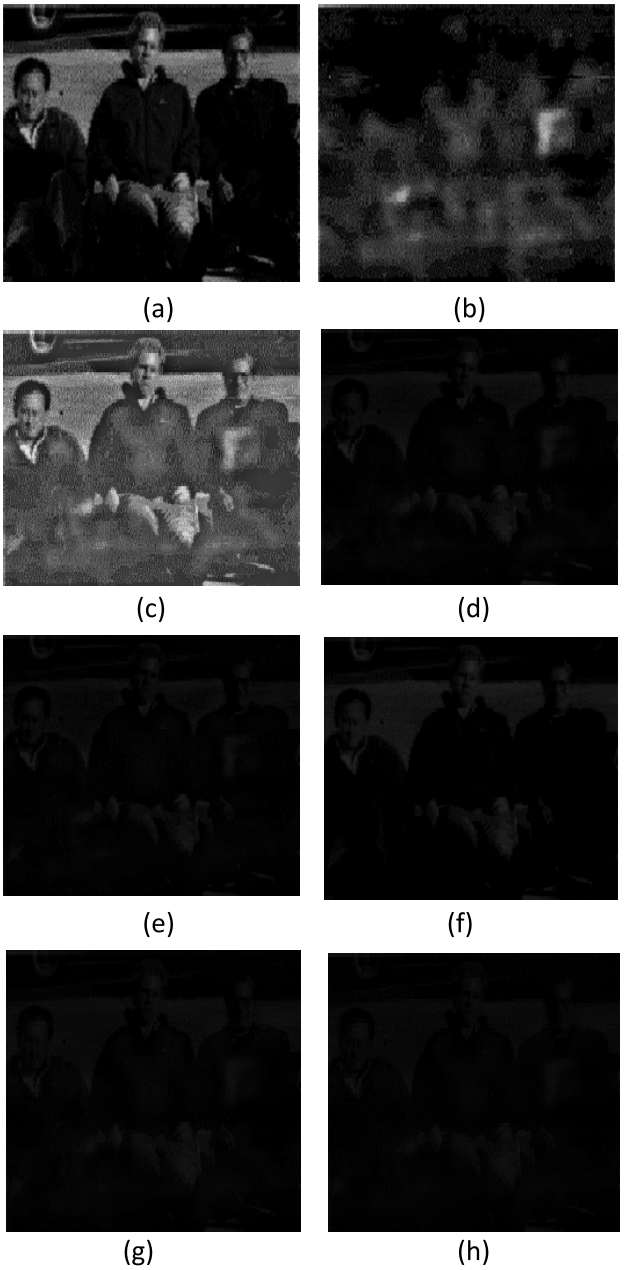

FIGURE 11. 3-man-car input images and fused images. (a) & (b) are input images; (c) proposed method; (d) MT-CT-PCA; (e) NSST-SF-PCNN; (f) median-average-DSWT-PCA; (g) DWT; (h) PCA.

MT-CT-PCA, NSST-SF-PCNN fusion, DWT and PCA are almost same; the brightness is not much bright; textures of a car are poor. The fusion effect of median-average method is even weaker and the background detail is not clear. In comparison to all above methods, the overall fusion performance of proposed method is better; the contrast is higher, texture information is much clear and the background detail is vivid.

In Figure 13, the fusion effect of PCA is worse and the overall impact is deficient. The texture details and background scenery for farmhouse are almost similar in NSST-SF-PCNN fusion, median-average and DWT but the contrast of DWT and median-average is blurred. From

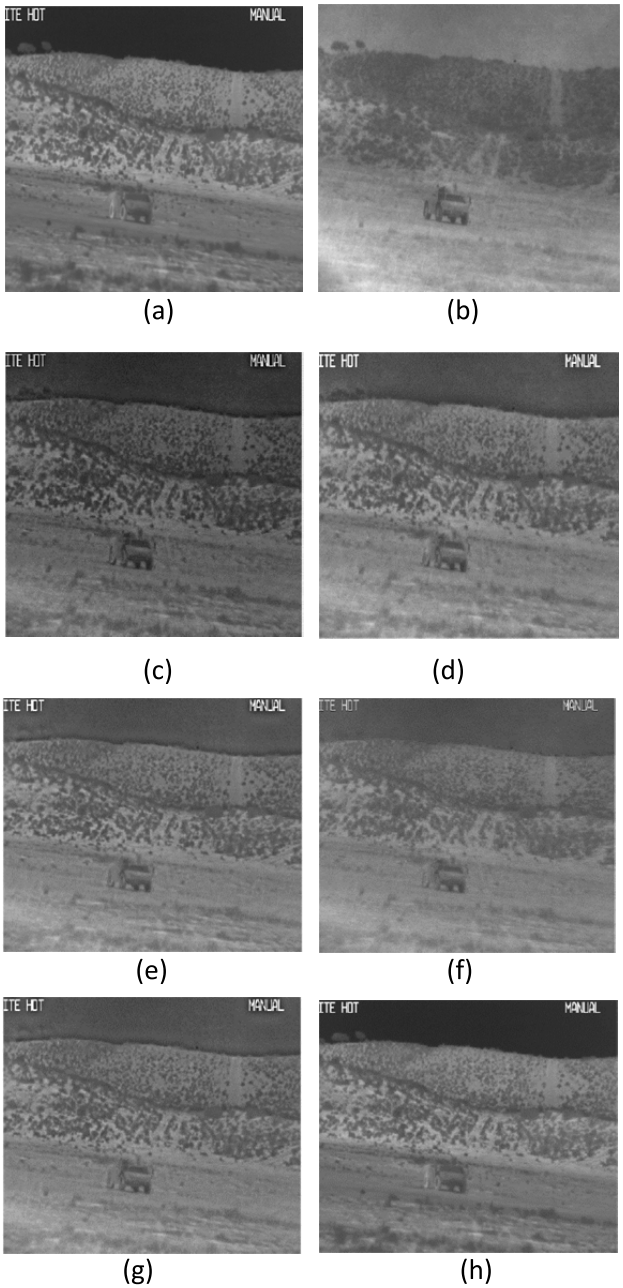

FIGURE 12. Desert-car input images and fused images. (a) & (b) are input images; (c) proposed method; (d) MT-CT-PCA; (e) NSST-SF-PCNN; (f) median-average-DSWT-PCA; (g) DWT; (h) PCA.

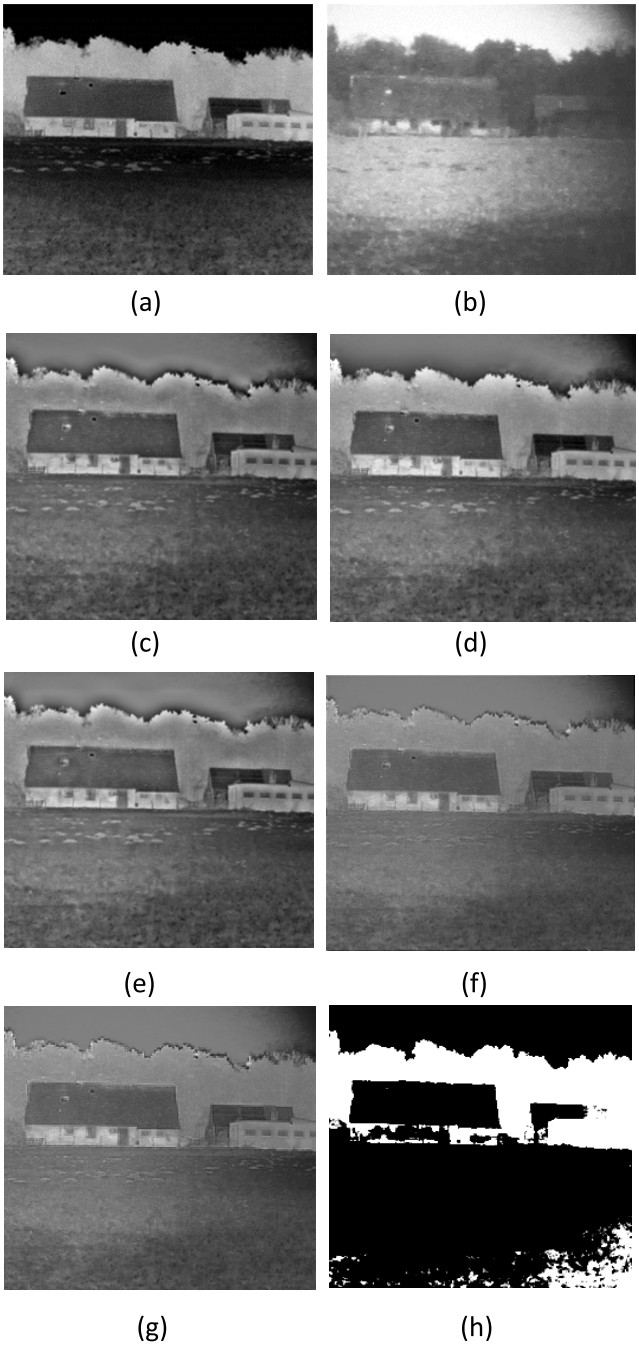

FIGURE 13. Farmhouse input images and fused images. (a) & (b) are input images; (c) proposed method; (d) MT-CT-PCA; (e) NSST-SF-PCNN; (f) median-average-DSWT-PCA; (g) DWT; (h) PCA.

Figure 13, it can be observed that the background scenery of MT-CT-PCA and proposed method are much clear and obvious than the rest of fusion schemes. Moreover, the proposed method has even more brightness, better contrast and high edge profile information than MT-CT-PCA that shows the superiority of our method than all aforementioned techniques.

The fusion effect of a bunker for IR and VI images is shown in Figure 14. The contrast of fused image for median-average and DWT is low in comparison to MT-CT-PCA and NSST-SF-PCNN methods but the textures of bunker are more apparent. The performance of PCA is inferior; the brightness is very high due to that, the overall effect of PCA is fuzzy. From Figure 14, it can be seen that the texture details of proposed method are much precise, and the information is more abundant in comparison to the conventional methods. Moreover, the overall effect of proposed method is best among the rest of the algorithms.

It can be seen from Figure 15 that the fusion results of book-man-clocks for MT-CT-PCA and NSST-SF-PCNN fusion methods are almost same, but their contrast is higher than median-average, DWT and PCA. In comparison to other techniques, the performance of DWT is inadequate for this

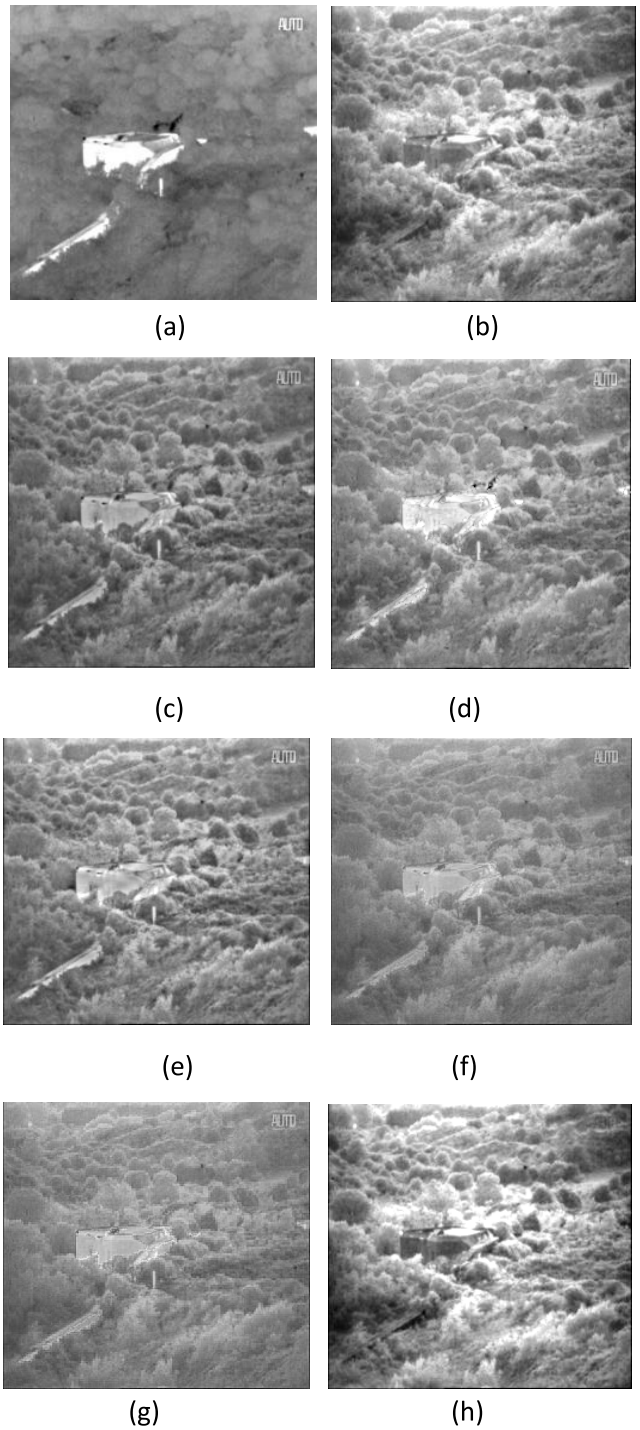

FIGURE 14. Bunker input images and fused images. (a) & (b) are input images; (c) proposed method; (d) MT-CT-PCA; (e) NSST-SF-PCNN; (f) median-average-DSWT-PCA; (g) DWT; (h) PCA.

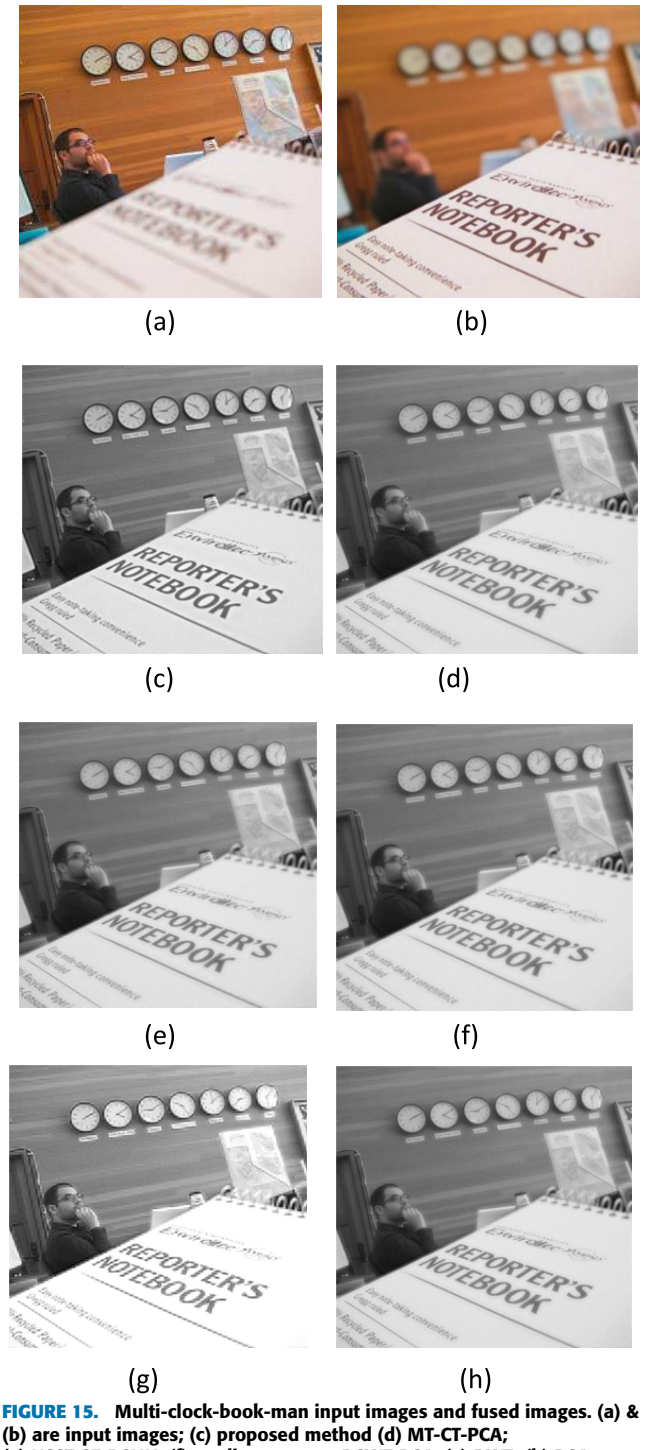

FIGURE 15. Multi-clock-book-man input images and fused images. (a) & (b) are input images; (c) proposed method (d) MT-CT-PCA; (e) NSST-SF-PCNN; (f) median-average-DSWT-PCA; (g) DWT; (h) PCA.

fused image. The contrast of proposed method is the best and the book letters are even more vivid than aforementioned methods. Additionally, the wall clocks are also obviously clear than other methods and fused image has sufficient information. Therefore, the overall quality of proposed method is much better, which shows superiority than other techniques.

It can be observed from Figure 16, the fusion effect of median-average, DWT and PCA look similar, the contrast and detail information are almost same and they produce satisfactory results. The proposed method, MT-CT-PCA and NSST-SF-PCNN fusion methods have much bright contrast and brightness than median-average, DWT and PCA. It can be seen in Figure 16 that the backside contrast and mountain in the proposed method are even more vivid than MT-CT-PCA and NSST-SF-PCNN fusion methods. Therefore, it can be analyzed that the overall performance of

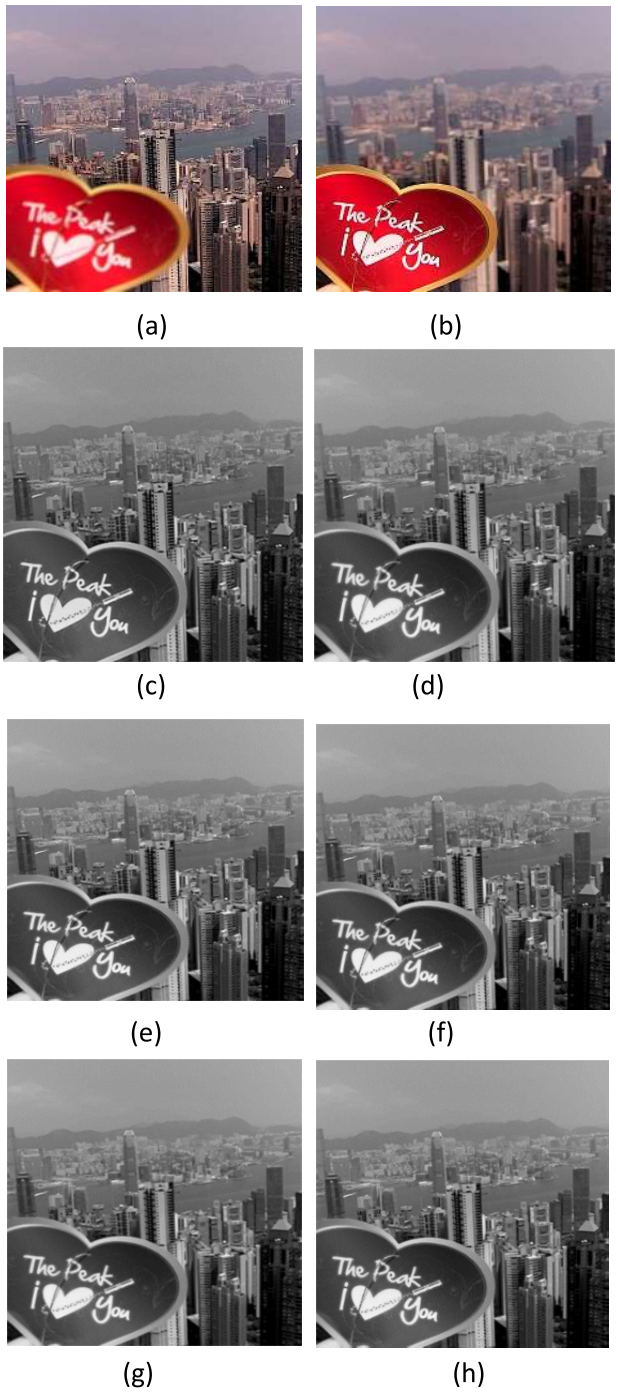

FIGURE 16. Multi-buildings input images and fused images. (a) & (b) are input images; (c) proposed method; (d) MT-CT-PCA; (e) NSST-SF-PCNN; (f) median-average-DSWT-PCA; (g) DWT; (h) PCA.

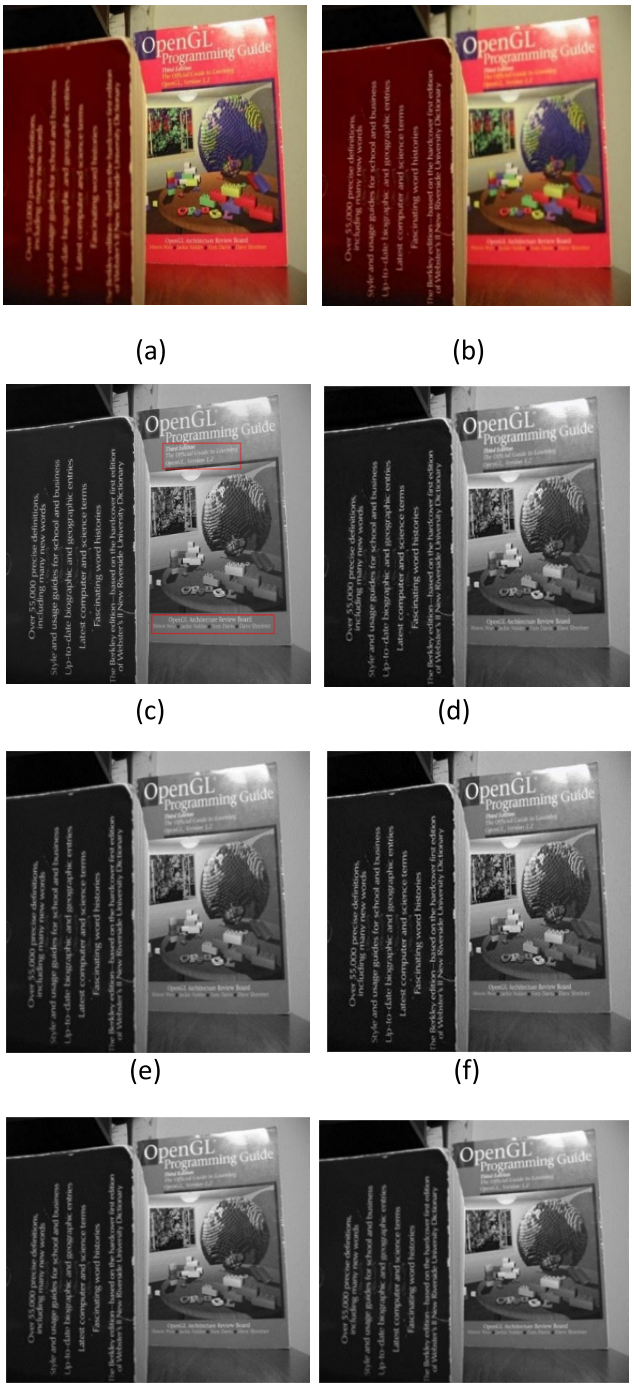

FIGURE 17. Multi-focus book input images and fused images. (a) & (b) are input images; (c) proposed method; (d) MT-CT-PCA; (e) NSST-SF-PCNN; (f) median-average-DSWT-PCA; (g) DWT; (h) PCA.

proposed method is higher than other conventional techniques, and it has high edge profile information with abundant information.

The MT-CT-PCA and NSST-SF-PCNN fusion schemes have better contrast, and letters are more clear in comparison to median-average, DWT and PCA, which is illustrated in Figure 17. The fusion effect of median-average, DWT and PCA look similar with a very slight difference; the letters on books are not bright and contrast is also reduced. The proposed method has much better results and the letters are more vivid than all aforementioned methods as depicted 17. The proposed method has less artifacts, more information in well-adjusted contrast. Besides, it preserves edge profile information, texture details and contours.

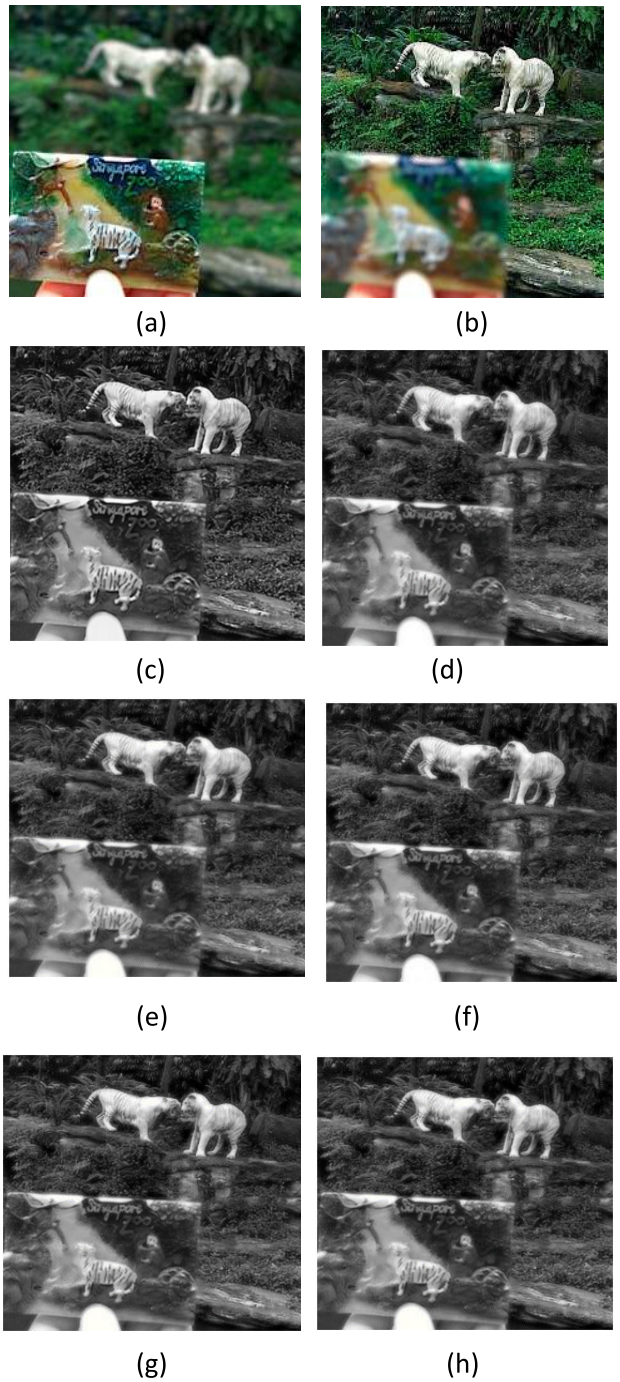

FIGURE 18. Multi-focus zoo-animal input images and fused images. (a) & (b) are input images; (c) proposed method; (d) MT-CT-PCA; (e) NSST-SF-PCNN; (f) median-average-DSWT-PCA; (g) DWT; (h) PCA.

The median-average method in Figure 18 for zoo-animal has more contrast than MT-CT-PCA, NSST-SF-PCNN, DWT and PCA techniques but the overall image quality is almost same. It can be analyzed that the proposed method for zoo-animal has better contrast and vivid information than other methods. It has also all detailed information from both input images with very less information loss and almost similar edge profile, texture and contour information as in both input images. Therefore, it is more obvious that proposed method is capable of fusing much information and other details from input images in comparison to other methods.

2) OBJECTIVE EVALUATION

In this paper, we also conduct objective analysis for the proposed method and state-of-art image fusion methods. This analysis is based on a mathematical calculation by some computable formulas. The selected objective evaluation assessments are given below:

a: AVERAGE PIXEL INTENSITY (API)

The API computes the index of fused contrast image and it is calculated by:

$$API = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} f(i,j)}{m,\, n} \tag{16}$$

$f(i,j)$ represents the pixel intensity at position (i,j) and $m \times n$ is the image size.

b: STANDARD DEVIATION (SD)

It computes the spread of information in fused image and it is computed by:

$$SD = \sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n} (f(i,j) - F)^2}{mn}} \tag{17}$$

Here $f(i,j)$ is the pixel value of image, F is the mean, and mn is the total size of image.

c: AVERAGE GRADIENT (AG)

It calculates the degree of sharpness and clarity in the image which is calculated by:

$$AG = \sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n} (f(i,j) - f(i+1,j))^2 + f(i,j) - f(i,j+1))^2}{mn}} \tag{18}$$

Here $f(i,j)$ is the pixel value of an image, mn is the number of rows and columns in an image.

d: SPATIAL FREQUENCY (SF)

The SF calculates the total information in a region of the image. It is obtained by:

$$SF = \sqrt{RF^2 + CF^2} \tag{19}$$

where RF and CF are expressed as:

$$RF = \sqrt{\frac{\sum_{i}\sum_{j} (f(i,j) - f(i,j-1))^2}{mn}} \tag{20}$$

$$CF = \sqrt{\frac{\sum_{i}\sum_{j} (f(i,j) - f(i-1,j))^2}{mn}} \tag{21}$$

TABLE 1. The quality assessment parameters comparison of proposed method and other state-of-art methods for RGB images visualized as gray images.

Measure	API	SD	AG	CC	SF	Time in Seconds
Input image Multi-modal medical image						
Proposed method	**69.564**	35.7018	13.0075	**0.7142**	**13.9485**	3.2191
MT-CT-PCA	53.5484	**36.0205**	**14.8901**	0.6213	12.8402	5.8018
NSST-SF-PCNN based	41.245	32.3508	9.4419	0.6828	12.6451	5.1094
median-average based DSWT-PCA	41.122	32.1220	8.2101	0.6910	11.9810	4.4109
DWT	40.54	32.2318	8.9922	0.6861	11.8905	3.0119
PCA	46.670	35.0126	7.5730	0.5429	12.0649	2.9910
Input image Multi-focus clock						
Proposed method						3.0187
MT-CT-PCA	**89.7575**	**48.2715**	9.1216	0.8709	**15.3765**	5.9101
NSST-SF-PCNN based	81.3710	48.1629	**9.2011**	0.8219	14.7510	6.0198
median-average based DSWT-PCA	79.2818	47.9905	8.9162	0.8012	14.8528	4.0156
DWT	81.0615	48.0170	9.0185	**0.8801**	14.6381	2.6791
PCA	72.1910	47.1520	8.1901	0.7991	12.1082	3.0167
Input image 3-Man						
Proposed method	**83.5635**	**38.1445**	**31.152**	**0.8014**	**21.1021**	4.2013
MT-CT-PCA	71.1083	31.1071	25.0817	0.6801	16.9012	6.1190
NSST-SF-PCNN based	69.0812	29.0829	24.1925	0.6011	15.1909	5.9185
median-average based DSWT-PCA	74.2205	31.3106	26.1108	0.7109	17.3164	5.0167
DWT	70.1018	31.0182	24.0901	0.6615	14.9910	3.6792
PCA	40.0192	18.0110	13.9172	0.2109	7.0312	2.9801
Input image desert-car						
Proposed method	**83.2065**	25.1021	**8.2620**	**0.8815**	**14.3315**	3.0169
MT-CT-PCA	76.0183	23.0151	7.1028	0.7919	12.1109	6.0916
NSST-SF-PCNN based	78.2017	**25.2429**	7.7501	0.8117	13.0116	5.1078
median-average based DSWT-PCA	73.8105	21.1083	6.6171	0.6911	11.0513	5.0109
DWT	75.1098	24.1006	7.3101	0.7210	12.0013	2.9017
PCA	75.9108	24.4109	7.5602	0.7402	12.2191	2.3018
Input image Farmhouse						
Proposed method	**89.1945**	**56.1417**	18.1971	**0.9081**	**22.1907**	4.4109
MT-CT-PCA	81.9102	51.0917	**19.1301**	0.8212	18.4510	7.0190
NSST-SF-PCNN based	72.6608	48.3170	16.0987	0.7218	17.0019	6.1092
median-average based DSWT-PCA	74.1402	50.0017	17.1092	0.7211	17.7810	5.9120
DWT	71.1390	46.2178	14.7801	0.6981	18.1420	3.3809
PCA	41.0368	32.0191	6.5810	0.3101	11.1901	3.1920
Input image Infrared and Visible image for Bunker						
Proposed method	**92.1109**	**29.9625**	**13.6170**	**0.9114**	17.0085	4.0107
MT-CT-PCA	78.2114	26.0187	10.1980	0.8410	14.8510	6.7126
NSST-SF-PCNN based	80.0258	27.8102	11.2180	0.8841	15.8180	6.0180
median-average based DSWT-PCA	76.6684	25.9810	13.0210	0.8114	**17.0919**	5.7801
DWT	75.2801	26.0010	12.0124	0.7919	14.7219	2.9096
PCA	72.0021	24.3202	10.0221	0.7409	15.5590	2.8070

e: CROSS-CORRELATION (CC)

This parameter is used to measure the similarity between the fused image and both input images. The CC is calculated as:

$$CC = \frac{S_{AF} + S_{BF}}{2} \tag{22}$$

$$S_{AF} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(a(i,j) - A)(f(i,j) - F)}{\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(a(i,j) - A)^2(\sum_{i=1}^{m}\sum_{j=1}^{n}(f(i,j) - F)^2}} \tag{23}$$

$$S_{BF} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(b(i,j) - B(f(i,j) - F)}{\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(b(i,j) - B)^2(\sum_{i=1}^{m}\sum_{j=1}^{n}(f(i,j) - F)^2}} \tag{24}$$

Here, $a(i,j)$, $b(i,j)$ represent the input image 1 and image 2, A and B are the mean of each. The value of CC is closer to 1 if the two images are more similar.

The values for quality assessment parameters are presented in Table 1 and Table 2. Table 1 presents the parameter comparison of RGB images from Figure 9-14 that are visualized as gray images. Table 2 presents the objective parameter comparison of RGB images for Figures 15-18 that are visualized as colored images. The quality assessment parameters used in this paper are API, SD, AG, CC and SF. Higher the value for API, SD, AG, CC and SF; better would be the quality of the fused image. In these Tables, bold values show the higher value for that quality parameter. It can be analyzed in Table 1 and Table 2; the proposed method also achieves best results than existing fusion techniques for quality parameters. However, the proposed method attains less value for some quality parameters but its overall performance is better that can be seen in Table 1 and 2. The fused image of proposed method has better contrast, more vivid, higher energy information, better edges, textures and smooth contours with minimal distortion and negligible information loss than the rest of fusion methods.

TABLE 2. The quality assessment parameters comparison of proposed method and other state-of-art methods for RGB images that visualized as colored images.

Measure	API	SD	AG	CC	SF	Time in seconds
Input image man-book						
Proposed method	**94.9410**	**66.1920**	**9.1591**	**0.9519**	**19.9782**	4.5701
MT-CT-PCA	86.1921	58.2940	7.2814	0.8619	15.5216	6.3183
NSST-SF-PCNN based	84.1019	57.9182	7.6519	0.8820	17.2150	5.8804
median-average based DSWT-PCA	78.8710	55.1298	7.0898	0.8410	16.1009	5.0127
DWT	72.1014	51.1092	6.4125	0.6680	11.0198	3.9809
PCA	79.9872	56.9879	7.2614	0.8119	14.1871	2.9014
Input image Peak building						
Proposed method	**91.3169**	**60.1025**	13.5766	**0.9204**	**24.0035**	4.9801
MT-CT-PCA	89.1502	55.2140	**13.5811**	0.8715	20.0087	7.0138
NSST-SF-PCNN based	85.3950	54.0129	11.0148	0.8117	21.0198	6.0916
median-average based DSWT-PCA	82.0018	52.0185	10.0291	0.7102	19.1940	5.9027
DWT	81.0901	53.8520	10.2986	0.7402	18.0182	4.0149
PCA	82.0015	52.0117	9.8201	0.7002	16.0184	3.8902
Input image Multi-focus book						
Proposed method	**81.4904**	**51.0184**	**17.6217**	**0.8801**	**29.0184**	5.1079
MT-CT-PCA	74.1015	46.0182	14.0192	0.8101	28.0143	7.2108
NSST-SF-PCNN based	71.9120	45.1025	15.0178	0.7994	26.9891	6.0905
median-average based DSWT-PCA	68.9104	42.9001	13.4130	0.7414	22.0182	6.1260
DWT	67.1001	40.1925	13.0096	0.6911	21.0185	4.8013
PCA	62.0907	36.0187	11.0192	0.7009	23.9871	3.3360
Input image zoo-animal						
Proposed method	**84.1490**	**58.7201**	20.8770	**0.9295**	**27.9761**	4.8601
MT-CT-PCA	75.0180	51.0198	15.1025	0.8116	24.1298	7.7203
NSST-SF-PCNN based	74.1001	48.2910	13.0158	0.8007	23.1025	5.8206
median-average based DSWT-PCA	77.0019	53.0187	**21.0109**	0.8202	20.0185	4.9902
DWT	69.0198	46.0182	12.0141	0.7132	17.1002	3.0915
PCA	67.1001	44.0185	14.9004	0.7515	19.2980	2.0156

3) COMPUTATION TIME COMPARISON

The computational efficiency of the proposed method and other state-of-art image fusion techniques are compared in Table 1 and Table 2. The experiments for computation time (t) are performed on MATLAB 2016b with 8 GB RAM and core i5 3.20 GHz CPU. It can be seen in the last column of Table 1 and Table 2 that the proposed method consumes very less time in comparison to MT-CT-PCA, NSST-SF-PCNN and median-average based DSWT-PCA. In addition, its qualitative and quantitative performance is also better than aforementioned state-of-art techniques. However, proposed method consumes little more time that DWT and PCA but the performance of proposed method is best that can be seen in subjective as well as objective evaluation parameters, which reveals the superiority of the proposed method.

V. CONCLUSION

In this research, we proposed a novel method that combines HEFGLG with NSCT and PCA-based conversion from multi-channel to a single-channel gray image. The experimental results substantiate that the proposed method combines the advantages of different sequential stages (HEFGLG, NSCT, improved fusion strategies, and gray-PCA). These combined methods provide better contrast, preserve sufficient energy information, achieve bright textures, smooth contours, and better visual effects than existing state-of-art techniques. Intuitively, the proposed method has tremendous fusion effects that can be observed from subjective as well as objective evaluation, which reveals the superiority of proposed work.

Many research points are still open and should be considered for future work. The next research direction would be to design more simple image fusion algorithm that plays a vital role in improving the image quality and applicable for real-time applications. Another research direction would be to use deep convolutional neural network for real-time monitoring applications of image fusion that consumes less time and produces high quality image.

REFERENCES

[1] A. Dogra, B. Goyal, and S. Agrawal, "From multi-scale decomposition to Non-Multi-Scale decomposition methods: A comprehensive survey of image fusion techniques and its applications," *IEEE Access*, vol. 5, pp. 16040–16067, 2017.

[2] X. Jin, Q. Jiang, S. Yao, D. Zhou, R. Nie, J. Hai, and K. He, "A survey of infrared and visual image fusion methods," *Infr. Phys. Technol.*, vol. 85, pp. 478–501, Sep. 2017.

[3] 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.

[4] T. Lianfang, J. Ahmed, D. Qiliang, B. Shankar, and S. Adnan, "Multi focus image fusion using combined median and average filter based hybrid stationary wavelet transform and principal component analysis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 6, pp. 34–41, 2018.

[5] J. Du, W. Li, and H. Tan, "Intrinsic image decomposition-based grey and pseudo-color medical image fusion," *IEEE Access*, vol. 7, pp. 56443–56456, 2019.

[6] X. Han, T. Lv, X. Song, T. Nie, H. Liang, B. He, and A. Kuijper, "An adaptive two-scale image fusion of visible and infrared images," *IEEE Access*, vol. 7, pp. 56341–56352, 2019.

[7] W. Li, L. Jia, and J. Du, "Multi-modal sensor medical image fusion based on multiple salient features with guided image filter," *IEEE Access*, vol. 7, pp. 173019–173033, 2019.

[8] K. C. Bhataria and B. K. Shah, "A review of image fusion techniques," in *Proc. 2nd Int. Conf. Comput. Methodol. Commun.*, Feb. 2018, pp. 114–123.

[9] M. H. Asmare, V. S. Asirvadam, L. Iznita, and A. F. M. Hani, "Image enhancement by fusion in contourlet transform," *Int. J. Electr. Eng. Inform.*, vol. 2, pp. 29–42, Nov. 2010.
[10] S. Liu, J. Wang, Y. Lu, S. Hu, X. Ma, and Y. Wu, "Multi-focus image fusion based on residual network in non-subsampled shearlet domain," *IEEE Access*, vol. 7, pp. 152043–152063, 2019.
[11] S. Liu, Y. Lu, J. Wang, S. Hu, J. Zhao, and Z. Zhu, "A new focus evaluation operator based on max-min filter and its application in high quality multi-focus image fusion," *Multidimensional Syst. Signal Process.*, vol. 31, pp. 569–590, Aug. 2019.
[12] S. Li, B. Yang, and J. Hu, "Performance comparison of different multi-resolution transforms for image fusion," *Inf. Fusion*, vol. 12, no. 2, pp. 74–84, Apr. 2011.
[13] S. Mane and S. D. Sawant, "Image fusion of CT/MRI using DWT, PCA methods and analog DSP processor," *Int. J. Eng. Res. Appl.*, vol. 4, no. 2, pp. 557–563, 2014.
[14] T. A. Soomro, T. M. Khan, M. A. U. Khan, J. Gao, M. Paul, and L. Zheng, "Impact of ICA-based image enhancement technique on retinal blood vessels segmentation," *IEEE Access*, vol. 6, pp. 3524–3538, 2018.
[15] T. A. Soomro, A. J. Afifi, A. Ali Shah, S. Soomro, G. A. Baloch, L. Zheng, M. Yin, and J. Gao, "Impact of image enhancement technique on CNN model for retinal blood vessels segmentation," *IEEE Access*, vol. 7, pp. 158183–158197, 2019.
[16] L. Zhan, Y. Zhuang, and L. Huang, "Infrared and visible images fusion method based on discrete wavelet transform," *J. Comput*, vol. 28, no. 2, pp. 57–71, 2017.
[17] J. Chen, X. Li, L. Luo, X. Mei, and J. Ma, "Infrared and visible image fusion based on target-enhanced multiscale transform decomposition," *Inf. Sci.*, vol. 508, pp. 64–78, Jan. 2020.
[18] M. Sharma, "A review: Image fusion techniques and applications," *Int. J. Comput. Sci. Inf. Technol.*, vol. 7, no. 3, pp. 1082–1085, 2016.
[19] L. Yang, B. L. Guo, and W. Ni, "Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform," *Neurocomputing*, vol. 72, nos. 1–3, pp. 203–211, Dec. 2008.
[20] P. Ashishgoud and G. Usha, "Image Fusion Using DWT & PCA," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 5, pp. 800–804, 2015.
[21] S. Liu, M. Shi, Z. Zhu, and J. Zhao, "Image fusion based on complex-shearlet domain with guided filtering," *Multidimensional Syst. Signal Process.*, vol. 28, no. 1, pp. 207–224, Jan. 2017.
[22] S. S. Kumar and S. Muttan, "PCA-based image fusion," *Proc. SPIE*, vol. 6233, May 2016, Art. no. 62331T.
[23] Y. Yang, "Multimodal medical image fusion through a new DWT based technique," in *Proc. 4th Int. Conf. Bioinf. Biomed. Eng.*, Jun. 2010, pp. 1–4.
[24] M. Qiguang and W. Baoshu, "A novel image fusion method using contourlet transform," in *Proc. Int. Conf. Commun., Circuits Syst.*, Jun. 2006, pp. 548–552.
[25] J. Adu, J. Gan, Y. Wang, and J. Huang, "Image fusion based on nonsubsampled contourlet transform for infrared and visible light image," *Infr. Phys. Technol.*, vol. 61, pp. 94–100, Nov. 2013.
[26] W. Kong, L. Zhang, and Y. Lei, "Novel fusion method for visible light and infrared images based on NSST–SF–PCNN," *Infr. Phys. Technol.*, vol. 65, pp. 103–112, Jul. 2014.
[27] Z. Qu, Y. Xing, and Y. Song, "An image enhancement method based on non-subsampled shearlet transform and directional information measurement," *Information*, vol. 9, no. 12, p. 308, Dec. 2018.
[28] H. Li, L. Liu, W. Huang, and C. Yue, "An improved fusion algorithm for infrared and visible images based on multi-scale transform," *Infr. Phys. Technol.*, vol. 74, pp. 28–37, Jan. 2016.
[29] X. Guo, R. Nie, J. Cao, D. Zhou, and W. Qian, "Fully convolutional network-based multifocus image fusion," *Neural Comput.*, vol. 30, no. 7, pp. 1775–1800, Jul. 2018.
[30] H. Li and X.-J. Wu, "Infrared and visible image fusion using latent low-rank representation," 2018, *arXiv:1804.08992*. [Online]. Available: http://arxiv.org/abs/1804.08992
[31] A. Ismail and E. Z. Fatma, "Image contrast enhancement techniques: A comparative study of performance," *Int. J. Comput. Appl.*, vol. 137, no. 13, pp. 43–48, Mar. 2016.
[32] M. Khan, M. Moatamedi, and B. Alzahabi, "A discriminative color-to-grayscale representation for retinal vessel segmentation," *Int. J. Multiphys.*, vol. 13, no. 1, 2019.
[33] A. L. Y. Meligy, "Modified fast gray level grouping approach for enhancing image contrast," *J. Theor. Appl. Inf. Technol.*, vol. 97, no. 21, 2019.
[34] A. L. Da Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, Oct. 2006.
[35] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.
[36] T. A. Soomro, M. A. U. Khan, J. Gao, T. M. Khan, and M. Paul, "Contrast normalization steps for increased sensitivity of a retinal image segmentation method," *Signal, Image Video Process.*, vol. 11, no. 8, pp. 1509–1517, Nov. 2017.
[37] R. I.-R. Bt, "Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios," Tech. Rep., 2011.
[38] L. Wang, B. Li, and L.-F. Tian, "EGGDD: An explicit dependency model for multi-modal medical image fusion in shift-invariant shearlet transform domain," *Inf. Fusion*, vol. 19, pp. 29–37, Sep. 2014.
[39] M. Kim, D. K. Han, and H. Ko, "Joint patch clustering-based dictionary learning for multimodal image fusion," *Inf. Fusion*, vol. 27, pp. 198–214, Jan. 2016.
[40] T. A. Soomro, A. J. Afifi, L. Zheng, S. Soomro, J. Gao, O. Hellwich, and M. Paul, "Deep learning models for retinal blood vessels segmentation: A review," *IEEE Access*, vol. 7, pp. 71696–71717, 2019.
[41] T. A. Soomro, A. J. Afifi, J. Gao, O. Hellwich, L. Zheng, and M. Paul, "Strided fully convolutional neural network for boosting the sensitivity of retinal blood vessels segmentation," *Expert Syst. Appl.*, vol. 134, pp. 36–52, Nov. 2019.
[42] H. Li, X.-J. Wu, and J. Kittler, "Infrared and visible image fusion using a deep learning framework," in *Proc. 24th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2018, pp. 2705–2710.
[43] Y. Liu, L. Dong, Y. Ji, and W. Xu, "Infrared and visible image fusion through details preservation," *Sensors*, vol. 19, no. 20, p. 4556, Oct. 2019.
[44] Z. Zhou, B. Wang, S. Li, and M. Dong, "Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters," *Inf. Fusion*, vol. 30, pp. 15–26, Jul. 2016.
[45] M. Nejati, S. Samavi, and S. Shirani, "Multi-focus image fusion using dictionary-based sparse representation," *Inf. Fusion*, vol. 25, pp. 72–84, Sep. 2015.
[46] Y. Liu, S. Liu, and Z. Wang, "Multi-focus image fusion with dense SIFT," *Inf. Fusion*, vol. 23, pp. 139–155, May 2015.
[47] L. Dong, Q. Yang, H. Wu, H. Xiao, and M. Xu, "High quality multi-spectral and panchromatic image fusion technologies based on curvelet transform," *Neurocomputing*, vol. 159, pp. 268–274, Jul. 2015.

JAMEEL AHMED BHUTTO received the bachelor of engineering degree in telecommunication engineering and the master of engineering degree in communication systems and networking from the Mehran University of Engineering and Technology, in 2010 and 2016, respectively. He is currently pursuing the Ph.D. degree in pattern recognition and intelligence system with the South China University of Technology, Guangzhou, China. His research interests include image fusion, image enhancement, artificial intelligence, machine vision, and machine learning.

TIAN LIANFANG (Senior Member, IEEE) was born in Jining, Shandong, China, in 1968. He received the B.S. and M.S. degrees in mechanical engineering from the University of Shandong, in 1991 and 1994, respectively, and the Ph.D. degree in electrical and mechanical engineering from the Harbin Institute of Technology, in 1997.

From 1997 to 2000, he was a Postdoctoral Researcher with the South China University of Technology, China. From 2000 to 2001, he was a Visiting Scholar with the University of Californian at Riverside, USA. From 2001 to 2004, he was a Visiting Scholar with the University of Pittsburgh, USA. Since 2003, he has been a Professor with the School of Automation, South China University of Technology, China. His research interests include intelligent robot, machine vision, machine learning, artificial intelligence, pattern recognition and intelligent systems, bioengineering, and so on.

QILIANG DU was born in Foshan, Guangdong, China, in 1980. He received the B.S. degree in automation and the Ph.D. degree in control theory and control engineering from the South China University of Technology, Guangzhou, China, in 2003 and 2008, respectively.

From 2009 to 2010, he was a Postdoctoral Researcher with the South China University of Technology. Since 2009, he has been a Vice Professor with the School of Automation Science and Engineering, South China University of Technology.

TOUFIQUE AHMED SOOMRO (Member, IEEE) received the B.E. degree in electronic engineering from the Mehran University of Engineering and Technology, Pakistan, in 2008, the M.Sc. degree in electrical and electronic engineering by research from University Technologi PETRONAS, Malaysia, in 2014, and the Ph.D. degree in AI and image processing from the School of Computing and Mathematics, Charles Sturt University, Australia, in 2018. He remained Research Assistant for six months in the School of Business Analytic in Cluster of Big Data Analysis, The University of Sydney, Australia. He is currently an Assistant Professor with the Department of Electronic Engineering, QUEST–Larkana, Pakistan. His research interests include most aspects of image enhancement methods, segmentation methods, classifications methods, and image analysis for medical images.

YU LUBIN was born in Hangzhou, Zhejiang, China, in 1994. He received the B.S. degree in automation from the South China University of Technology, Guangzhou, China, in 2016, where he is currently pursuing the Ph.D. degree.

MUHAMMAD FAIZAN TAHIR received the B.Sc. degree in electrical engineering from the University of Engineering and Technology Taxila, Taxila, Pakistan, in 2011, the M.S. degree in electrical engineering from The University of Lahore, Lahore, Pakistan, in 2015, and the Ph.D. degree in power system and automation from the South China University of Technology, Guangzhou, China, in 2020. From 2012 to 2016, he was a Lecturer with The University of Lahore. He is currently a Postdoctoral Fellow with the South China University of Technology. His research interests include renewable energy integration, demand response, load management, power system optimization, and integrated energy system planning.