ISSN: 0067-2904

Image Fusion Based on the Biorthogonal Wavelet Transform and Average Gradient

Iman M.G. Alwan
Department of Computer, College of Education for Women, University of Baghdad, Baghdad, Iraq

Received: 29/1/2023 Accepted: 9/9/2023 Published: 30/11/2024

Abstract

The goal of fusing multi-focus images is to obtain one image that has all the significant features from each input. The fusion process is needed because of the limitations of the optical lens depth of field that is used to capture images, so images of various focused regions are produced. In this paper, a multi-focus image fusion algorithm is proposed. It is based on utilizing the biorthogonal wavelet transform to extract details and edges from the input images by making the approximation subband equal zero and applying an inverse transform to get images that have only edges, lines, and details. The average gradient metric, which is used to represent sharpness and clarity, is calculated as an activity measurement for each NxN block of the resulted edge images and is used to merge the corresponding blocks of the multi-focus input images to produce the fused one. Consistency verification is used to improve the fusion process. The performance of the proposed method was evaluated and contrasted with a number of other state-of-the-art fusion methods. Experimental results clearly show that the suggested approach is a feasible and efficient multi-focus imaging technique.

.

Keywords: Multi-focus image fusion, Biorthogonal wavelet transform, Average gradient, Consistency verification.

دمج الصور بالاعتماد على تحويلة المويجة المتعامدة و متوسط الانحدار

إيمان محمد جعفر علوان

قسم الحاسوب, كلية التربية للبنات, جامعة بغداد, بغداد, العراق

الخلاصة

الهدف من دمج الصور متعددة التركيز البؤري هو الحصول على صورة واحدة تحتوي على جميع الميزات المهمة من كل إدخال. هناك حاجة إلى عملية دمج الصور متعددة التركيز بسبب محدودية عمق مجال العدسات البصرية المستعملة لالتقاط الصور ، لذلك يتم إنتاج صور لمناطق مختلفة في التركيز. في هذا البحث، تم اقتراح خوارزمية دمج الصور متعددة التركيز بالاعتماد على استعمال التحويل المويجي المتعامد (Biorthogonal Wavelet Transform) لاستخراج تفاصيل وحواف الصور المدخلة عن طريق جعل النطاق الفرعي التقريبي (Low–Low subband) يساوي الصفر وتطبيق التحويل العكسي للمويجة (Inverse Biorthogonal Wavelet Transform) للحصول على صور ذات حواف وخطوط وتفاصيل فقط. تقسم هذه

*Email: ainms_66@coeduw.uobaghdad.edu.iq

الصور الى كتل بحجم NxNو احتساب متوسط التدرج اللوني (Average Gradient) لكل كتلة، والذي يستعمل لتمثيل الحدة والوضوح، كمقياس لاختيار الكتلة (block) ذات التركيز الاعلى للصور المدخلة. تم استعمال مبدأ التحقق من الاتساق في تحسين عملية دمج الصور. تم تقييم أداء الطريقة المقترحة ومقارنتها مع عدد من أحدث طرق الاندماج الحالية. تظهر النتائج التجريبية بوضوح أن النهج المقترح فعال جدا.

1. Introduction

The fusion of images is utilized to collect necessary and particular information from different input images into a single one. The resulted image should contain all relevant information transferred from the input image set. In a network of visual sensors, the camera sensor usually captures images with a blurring effect in some areas, so to get one image of all objects in focus, a fusion technique is used to make the resulted image more convenient for recognition, detection missions, and visualization [1]. Fusion techniques are implemented in four major sets: hybrid techniques, techniques based on the spatial domain and transformation domain, and deep learning-based techniques, which have gained a lot of popularity recently. Combining techniques that deal specifically with pixels and operations based on pixel density are known as "spatial domain approaches," whereas fusion of the transform domain involves image transformation by utilizing mathematical tools. Three distinct approaches to the spatial domain are pixel-based, block-based, and region-based. These techniques seek to choose pixels, blocks, or regions with more information. In pixel-based approaches, a sharpness analysis is carried out for every individual pixel in the image, and a fusion image is produced by choosing the pixels that are thought to be clear. In block-based techniques, the input images are divided into defined-size, tile-like portions. This method is used to calculate the activity level for each block [2]. Numerous strategies have been laid out in the literature for the spatial and transformation domains. The biorthogonal wavelet transform has a linear phase property, so it reduces image edge distortion. With filters of finite impulse response, typical image reconstruction can be obtained [3]. The proposed method depends on the transform domain and the spatial domain. It uses the biorthogonal wavelet transform to get edges and details from multi-focus images by setting the approximation subband to zero and then using the inverse transform to get rid of the coarse information and keep the fine information. The resulting images have only edge information. To detect focus regions, the resulted images are divided into non-overlapping N×N blocks, and then the average gradient (AG) metric is calculated for each block. Since AG reflects the sharpness and clarity of every block, blocks with higher AG values mean focused blocks, while blocks with lower AG values mean defocused blocks. According to this activity level measurement, blocks of the corresponding multi-focus input images are fused. To refine the fusion process and inhibit the blocking effect in the output image, consistency verification, which was proposed in [4], is adopted.

The contribution of this paper can be condensed as follows:

(1) A block-based image fusion algorithm is presented. It is based on the use of the biorthogonal wavelet transform to extract edge information from the input images. This technique gives perfect edge representation, hence perfect focus area detection.

(2) The average gradient metric (AG) is used as a block-focus measurement in the spatial domain, which realizes the more informative regions.

(3) Consistency verification is utilized to prevent inappropriate block choice and, hence, blocking effect.

The following sections are organized as follows: Section 2 demonstrates the related works; Section 3 explains the biorthogonal wavelet transform; Section 4 explains the proposed fusion

algorithm; Section 4 states the experimental results; and finally, Section 5 states the conclusion.

2. Related Works

In the literature, many methodologies for multi-focus image fusion have been offered. In [5], a new pixel-level autofocus image fusion algorithm is presented. This approach focuses on non-homogeneous input picture regions with precise features. Local standard deviation as a criterion is utilized to accomplish the fusion process. The block from the two input multi-focus images with a higher standard deviation is chosen to construct the fused one. This process is executed for the entire input image. Contrast and correlation criteria have been used to evaluate the fusion process. In [6], a fusion technique combining the proposed focusing filter and the Dual Tree-Complex Wavelet Transform is provided to improve the focus of the fused image and produce the highest quality image. Two filters—a Wiener filter and a sharpening filter—combine to form the focusing filter. Prior to utilizing the dual tree-complex wavelet transform for the fusion operation, this filter is applied. The fused image was created using the average-fusion rule and maximum-fusion rule, which are typical fusion rules. The suggested fusion algorithm's performance was compared to that of the current algorithms in the experiment employing the focus operators. The outcomes demonstrated that the proposed method is superior to these fusion methods in terms of quality. An additional image fusion method at the pixel level includes two phases: first, locating the source image's edges; then, a genetic algorithm that depends on these inputs finds weights for the fusion process. Different quality evaluation criteria have been used, which show superior results with respect to the wavelet transform-based algorithm and a reduction in run time [7].

In [8], a multi-focus image fusion method depends on using a guided image filter to obtain base and detail layers, and various methods are used to find focus regions. In the transform domain, the input images are decomposed into subbands, and the fusion coefficients come after applying some particular fusion rules. The fused image is then produced by applying the inverse transform. Experiments with 325 images reveal very good results in terms of clarity and robustness to noise compared to other traditional methods. The most commonly used transform-based algorithms include the wavelet transform [9]. In [10], while using DWT to deconstruct the source images, two distinct window-based fusion algorithms are used to individually combine the low- and high-frequency components. The method picks the low-frequency coefficients with the highest sharpness focus measure as the coefficients for the fused image. For the high-frequency sub-band coefficients, a maximum energy-based fusion scheme is suggested. The combined coefficients are subjected to a consistency testing technique in order to ensure the homogeneity of the final fused image. The proposed algorithm produces superior visual quality and objective assessment indices than pixel averaging, FSD pyramid, and DWT methods. In [11], the contourlet transform was used. The fusion algorithm and a denoising algorithm are coupled in the suggested work to counteract the effects of noise. In the suggested strategy, a level-dependent threshold based on the mean and median of the absolute contourlet coefficients was employed, as well as the standard deviation of the contourlet coefficients. It also depends on the degree to which contourlet coefficients are decomposed. It performs better than thresholds because noise has a varying impact on coefficients at different levels. As a result, the proposed threshold performs denoising better than existing thresholds. To show how well this proposed method works, it has been compared to the discrete wavelet transform, DTCWT, MWT, SWT, LWT, PCA, LP, and sharp-fusion-based methods. This proposed method outperforms other state-of-the-art methodologies, according to both qualitative and quantitative results.

In [12], multi-scale geometric analysis using the curvelet transform, which has anisotropy, was used for multi-focus image fusion, and fusion rules that work well for multi-focus image fusion were used. The algorithm was compared with the Laplacian pyramid transform and the wavelet transform. The fusion images of the wavelet transform and the curvelet transform have good results. Subtle ghosting may be seen in the wavelet-transformed merged image, while this phenomenon was discarded in the proposed algorithm. In [13], the image fusion technique relied on the discrete cosine transform and spatial frequency. In this algorithm, DCT separates the high-frequency components and treats the spatial frequency with the variation of pixel locality. The resulting image lacked any blurring or blocking artifacts. It performed better than conventional DCT, SWT, and many other cutting-edge methods. In [9], a traditional MIF system based on quarter-shift dual-tree complex wavelet transform (qshiftN DTCWT) and modified principal component analysis (MPCA) in the Laplacian pyramid (LP) domain is proposed to get the focused picture from a large number of source images. Since the DWT of image signals generates a nonredundant representation of them, a good spectral and spatial localization of the transformed image can be obtained, as can the low computational complexity with respect to other extended multi-resolution transforms.

In [14], a unique conditional random field optimization (mf-CRF) model is presented. When given two input images, the mf-CRF methodology achieves a globally optimal solution while preserving the benefits of pixel-domain and multi-spectral-domain fusion algorithms by tackling the proposed energy minimization problem. According to the experimental data that has been given, the suggested mf-CRF model outperforms modern image fusion techniques. To demonstrate the generalizability of the method, applications of the mf-CRF model to visible-infrared image fusion and medical picture fusion are also demonstrated. In [15], a decision map-based fusion method is used to improve the quality and clarity of the combined images. This method is based on a multi-focus image fusion algorithm that uses feedback mechanisms and a network structure that combines CNN and a transformer. CNN acquires local features, while Transformer gets global features, respectively. This proposed system improves the ability to extract features and retain information more effectively. There was a feedback mechanism in the process of feature extraction in order to iteratively get more contextual information and enhance the reliability of the fusion process. This study compares the technique in this paper to seven other sophisticated fusion methods using color and grayscale datasets.

The results demonstrate that this approach is superior in both subjective and objective assessments. In [16], the author suggested an integrated strategy for multi-focus image fusion analysis for the fused image coefficient selection process using DWT and computer vision. He thoroughly examined and improved the existing wavelet transform methods and multi-focus image fusion guidelines for extracting object information. The wavelet transform makes use of real localization segments, and computer vision has sped up the processing of combined images while analyzing object focus in high-frequency steps. In iterative trials, the wavelet basis function and wavelet decomposition level are used for image fusion to get high-quality information from the two pictures that have been put together. The wavelet transformation of the high-frequency coefficients makes the fusion picture features in the regional contrast and frequency domains of the object more reliable. In [17], the biorthogonal wavelet transform is used to break up the source images. The absolute maximum rule is used

to combine the low-frequency subbands of the input images, and the regional energy-weighted method is used to combine the coefficients of the high-frequency subbands.

3. The Biorthogonal Wavelet Transform

The biorthogonal wavelet transforms act like orthogonal wavelets, and they are also more flexible. Their filter banks offer ideal analysis and reconstruction signals. Discrete wavelet transforms have basis functions of single orthogonality, while biorthogonal wavelet transforms have a pair of dual biorthogonal basis functions for analysis and synthesis steps, respectively. The decomposition of an image using DWT is implemented by passing it through a low-pass filter and a high-pass filter to get the approximation and detail subbands. The outputs are down-sampled by 2. In the case of biorthogonal, there are two scaling functions { , , $\boldsymbol{\phi}$}. that may generate different multiresolution analyses, and accordingly, two different wavelet functions {$\boldsymbol{\psi}$, ,$\boldsymbol{\psi}$.}. So the numbers M and N of coefficients in the scaling sequences {**a**, ,.} may differ. The scaling sequences must satisfy the following biorthogonality condition.[18]

$$\sum_{n\in Z} a_n \tilde{a}_{n+2m} = 2 \cdot \delta_{m,0} \quad (1)$$

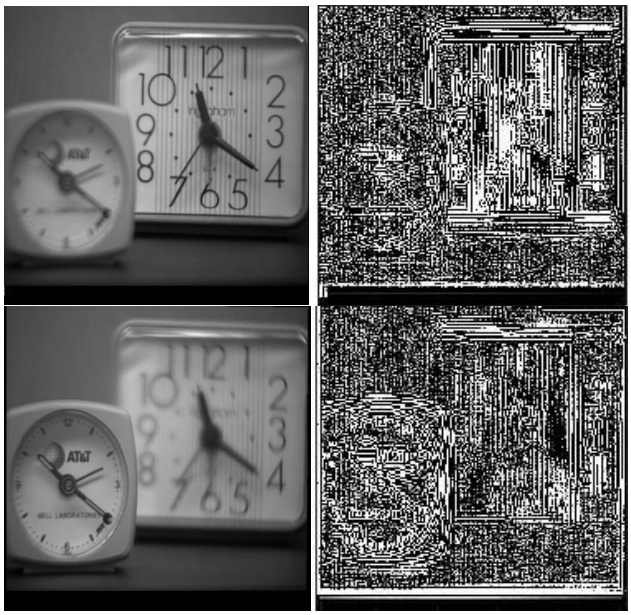
So, the sequences of wavelets are determined as

$$b_n = (-1)^n \tilde{a}_{M-1-n} \qquad (n=0, \ldots, N-1) \quad (2)$$

$$\tilde{b}_n = (-1)^n a_{M-1-n} \qquad (n= 0,\ldots,N-1) \quad (3)$$

4. The Proposed Fusion Algorithm

The biorthogonal wavelet transform is used to find areas of focus in the source images, which are sharper and clearer than areas that aren't in focus. This is done by extracting edges and details from the source images. This can be done by making the low-frequency sub-band, which represents the smoothed and subsampled version of the original image, equal to zero and keeping the detail subbands, then applying the inverse biorthogonal wavelet transform. Figure 1 exhibits the multi-focus input images "Clock" and the detail images that were extracted from them.



(a) Input Images **(b) Detail Image**

Figure 1: Input images "Clock" and detail images resulted from using BWT

The proposed block-based image fusion algorithm relies on efficiently using the detail images to distinguish the focused regions from the defocused regions by using the average gradient metric (AG), which measures the sharpness and clarity of the image. The detail images are divided into NxN non-overlapping blocks, and the AG metric is calculated for each block. Blocks with a higher value of AG refer to sharper blocks of the input multi-focus images. For proper fusion, if multiple blocks of focused areas derive from the first input source image, the blocks of all the areas should be picked from the first one. As a result of undesirable noise, several blocks may be chosen from the second input image. This causes a blurring effect in the fused image. To override this problem, consistency verification is used. This method was proposed by Li et al. [4]. It is executed by employing the majority filter to reform the chosen block in such a manner that if a block is from the first image and the encirclement blocks are from the second image, the algorithm will pick the picked block from the majority blocks. The resultant fused image is complete in sharpness and clarity because the fused image consists of the focused blocks of the source input images.

In this paper, the input multi-focus images are supposed to be registered. The following steps elucidate the steps of the proposed algorithm for two input source images I_X, I_Y:

Step 1. Decompose the input source images I_X, I_Y, into low frequency subbands (coarser resolution level) and high frequency subbands, make the low frequency band of both transformed images equal to zeros, then apply the inverse biorthogonal wavelet transform to get detail images D_X, D_Y .

Step 2. Divide the detail images, D_X, D_Y into non overlapping blocks of size $N \times N$, apply average gradient metric (AG) for each block.

Step 3. Depending on the value of the average gradient of each block, which is considered an activity level measurement, a binary decision matrix B^{AG} is constructed.

Step 4. The binary decision matrix B^{AG}, is used in window-based consistency verification mode to generate a decision matrix M^{AG}.

Step 5. The fused image F_{XY} is constructed by choosing the corresponding $N \times N$ block of the input images I_X, I_Y depending on the decision matrix M^{AG} values.

Figure 2 shows the general framework of the proposed system.

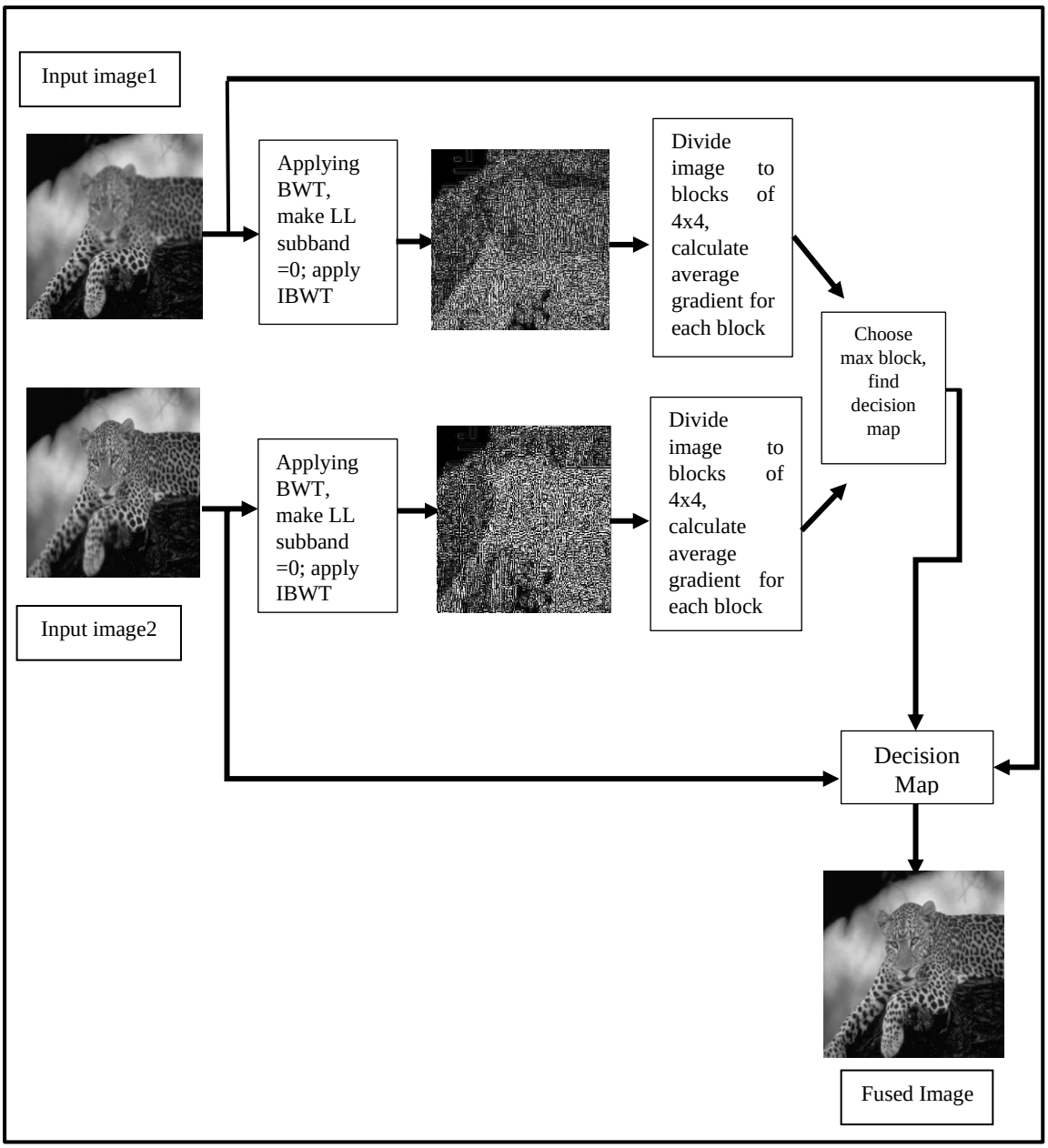


Figure 2: The structure of the proposed fusion system

4.1 The Fusion Rule

In the fusion process, the key issue to be considered is how to measure the activity level of the detail images, which indicate the sharpness and clarity of the input source images. In this proposed algorithm, the detail images are divided into non-overlapping blocks of size $N \times N$. Average gradient metric (AG) is used as a measure of sharpness for each block. Eq. (4) represents it [19].

$$AG = \frac{\sum_i \sum_j ((im_{i,j}-im_{i+1,j})^2+(im_{i,j}-im_{i,j+1})^2)^{\frac{1}{2}}}{mxn} \quad (4)$$

where, $im_{i,j}$ refers to the pixel at position (i,j) in each block of the detail images D_X, D_Y, m,n represent the block size. In this paper, the block size is 4×4.

Let $AG_{(i,j)}^{D_x}$, $AG_{(i,j)}^{D_y}$ denote the average gradient of each block of detail images extracted from input multi-focus images respectively. Then, $AG_{(i,j)}^{D_x}$ and $AG_{(i,j)}^{D_y}$ are compared to decide

which block is more probable to be categorized as the focused block. The binary decisionmatrix B^{AG} is constructed as follows:

$$B_{(i,j)}^{AG} = \begin{cases} 1 & AG_{(i,j)}^{D_x} \geq AG_{(i,j)}^{D_y} \\ -1 & otherwise \end{cases} \quad (5)$$

$$W_m(i,j) = \frac{1}{25}\sum_{r=-2}^{2}\sum_{l=-2}^{2} B^{AG}(i+r, j+l) \quad (6)$$

where, $W_m(i,j)$ represents the majority filter, B^{AG} is the Binary decision matrix
The fused image is found as:

$$F_{AB}(m,n) = \begin{cases} I_A(m,n) & W_m(i,j) > 0 \\ I_B(m,n) & W_m(i,j) \leq 0 \end{cases} \quad (7)$$

where, F_{AB} is the fused image, I_A, I_B are the input multi-focus images.

5. Experimental Results:

The experimental results of the proposed algorithm are explained in this section, as is the evaluation of this algorithm. Actually, evaluation of image fusion algorithms is labeled into two types: one requires a reference image, whereas the other does not. In this paper, the two methods are followed.

5.1 Performance Measures

Various parameters are used in the evaluation of fusion algorithms when no reference images are available [20], [14]. In this research, four criteria of quantitative measurement are used, where maximal values indicate the best performance. These criteria are:

- Mutual Information (MI): indicates the amount of information content that is correlated in the fused image with respect to the source input ones.

$$MI = MI_{XF} + MI_{YF} \quad (8)$$

where, MI_{XF} , MI_{YF} represent the mutual information between source image X, source image Y, and the fused image F, respectively.

$$MI_{XF} = \sum_x \sum_f p_{X,F}(x,f) log_2 \frac{p_{X,F}(x,f)}{p_x(x)P_F(f)} \quad (9)$$

$$MI_{YF} = \sum_y \sum_f P_{y,F}(y,f) log_2 \frac{P_{Y,F}(y,f)}{P_Y(y)P_F(f)} \quad (10)$$

- Petrovic Metric Parameter ($Q^{XY/F}$) is used to calculate how much information is transferred from input source images (X, Y) to fused one (F). The value of $(Q^{XY/F})$ equals to one means fusion without any information loss, so a closed value of $(Q^{XY/F})$ to one means a precise fusion result. It is proposed by Xydeas and Petrovic [21].
- Visual Information Fidelity for Fusion (VIFF) metric that measures an image's visual information [22].
- Spatial Frequency (SF), which indicates the level of activity of the image F (level of overall information) [20].

$$SF = \sqrt{RF^2 + CF^2} \quad (11)$$

where,

$$RF = \sqrt{\frac{\sum_i \sum_j (f(i,j) - f(i,j-1)^2}{m \times n}} \quad (12)$$

$$CF = \sqrt{\frac{\sum_i \sum_j (f(i,j) - f(i-1,j)^2}{m \times n}} \quad (13)$$

The parameters used with reference images are the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM). These are:

- Peak Signal to Noise Ratio (PSNR): the PSNR is usually used to evaluate the fused image quality [23]. Greater value determines the best fusion; it is set by Eq. (14):

$$PSNR = 10log_{10}(\frac{L^2}{\frac{1}{M\times N}\sum_{m=1}^{M}\sum_{n=1}^{N}[R(m,n)-F(m,n)]^2}) \quad (14)$$

where L is the maximum possible value of the image's pixels.

- SSIM: It is used for comparison between the fused and source images in terms of local patterns for pixel intensities [24]. It is defined by Eq. (15):

$$SSIM(R,F) = \frac{2\mu_R\mu_F+C_1}{\mu_R^2+\mu_F^2+C_1} \times \frac{2\sigma_{RF}+C_2}{\sigma_R^2+\sigma_F^2+C_2} \quad (15)$$

5.2 Experiment Analysis

The evaluation of the proposed algorithm consists of two parts. The first part has been implemented by using several test gray-scale multi-focus input images of size 256 × 256 without the availability of reference ones. The data set is available at https://www.mathworks.com/matlabcentral/fileexchange/70109-multi-focus-image-fusion-dataset. Generally, the input source images are supposed to be registered. The second part has been executed on multi-focus images generated artificially by applying an average filter of 5 × 5 window for both the left and right sides of them to apply the PSNR and SSIM metrics for fused and reference ones. Four images (numbers, a clock, a house, and peppers) were used. The simulation of the proposed algorithm has been executed using MATLAB R2016a with an Intel Core i7 2.7 GHz processor and 4 GB of RAM. Figure 2-a shows six of these non-referenced input source images, and Figure 2-b shows the resulted fused images. Figure 3 displays the test multi-focus images that were generated artificially and the fused ones that resulted, where one can notice that all the details and edges of the input sources are transferred to the fused one with no blurring effect.

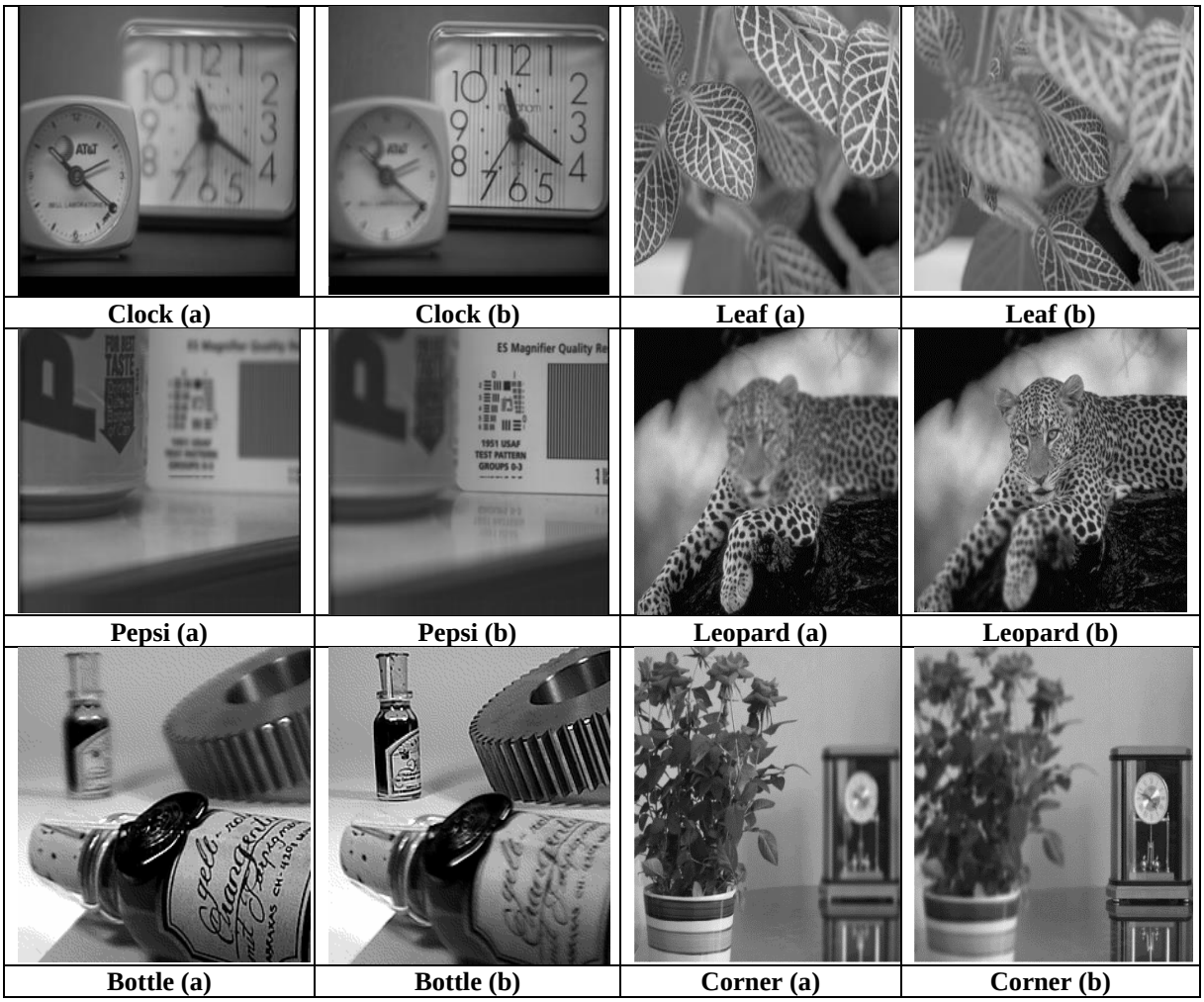


Figure 2 a: Test multi-focus input images

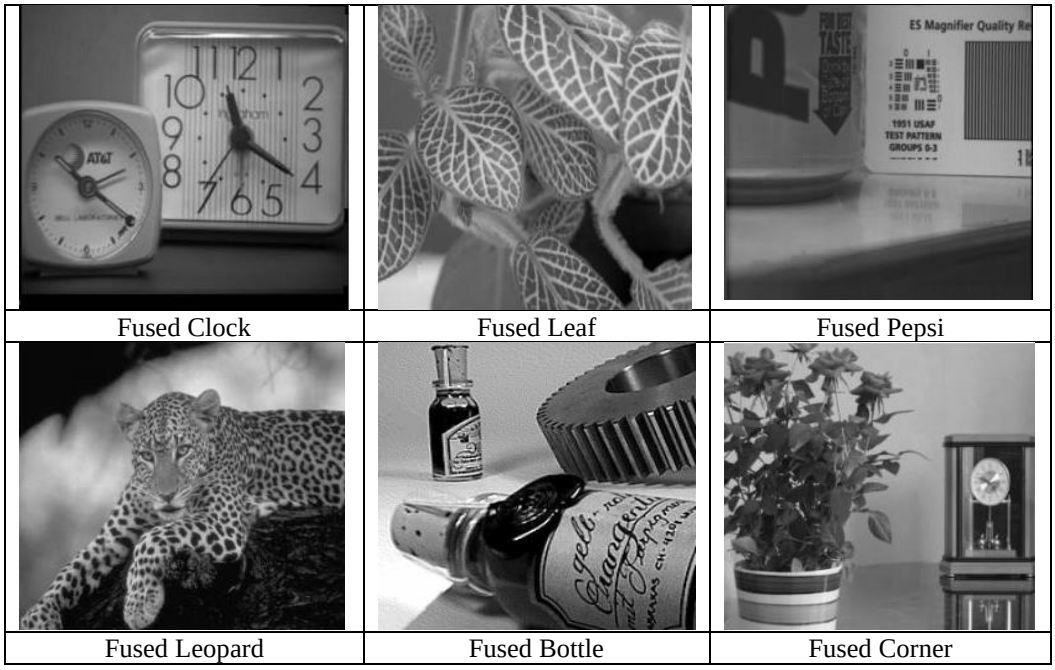


Figure 2 b: Resulting fused multi-focus test images of the proposed method

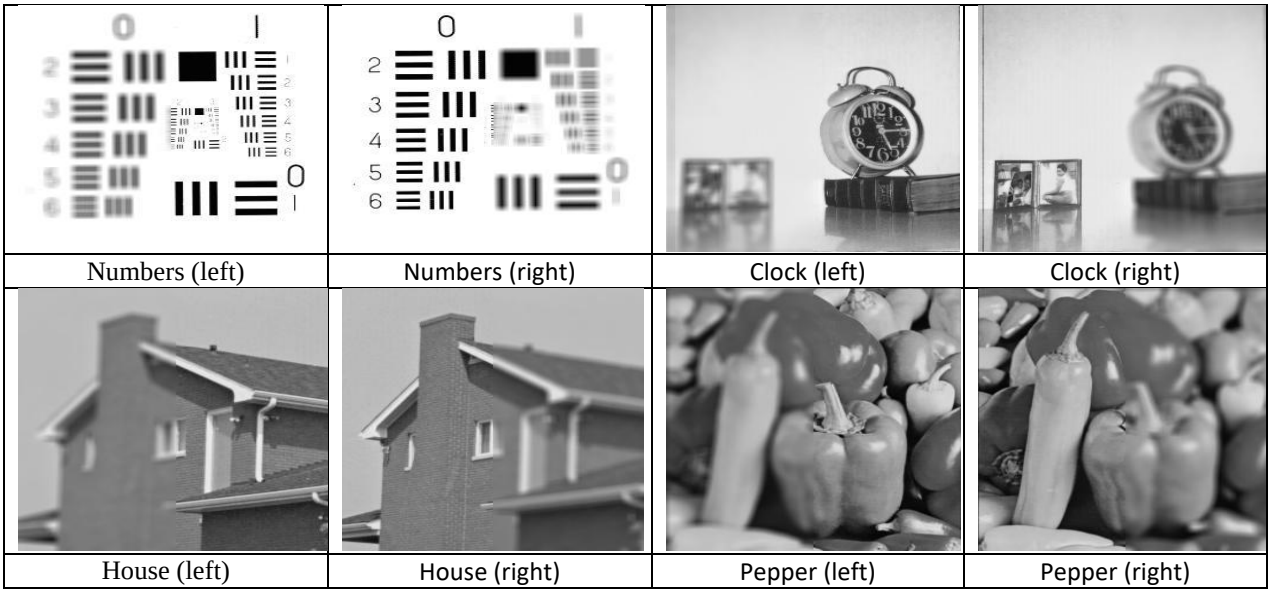


Figure 3: Test Images (blurred artificially)

The experiments revealed that the proposed methodology produces fine edges, high visual quality, high lucidity of objects, and less degradation. The proposed fusion algorithm produces high-quality image outputs. To elucidate the superiority of the proposed algorithm, performance evaluation criteria (MI, $Q^{XY/F}$, VIFF, SF) for non-reference images and (PSNR, SSIM) for reference images are applied. The obtained results are compared with some state-of-the-art algorithms that are transform-based and block-based. They are: DWT with consistency verification (DWTCV) [4], and multi-scale guided filter image fusion (MGFF). This work introduces a new visual saliency detection method based on a guided image filter. This filter can extract relevant sections from visually disparate images of the same scene [25]. Multi-scale image fusion is based on the calculation of variance in the discrete cosine transform domain (DCTVAR) [26], and finally, in [27], principle components are calculated

for the multi-scale coefficients of the wavelet transform domain, and weights for the fusion rule will be determined by averaging the principal components of all these pertinent decomposed elements (DWTPCAv). Also, the proposed algorithm has been compared with the algorithm in [28] (saliency), where image fusion is based on visual saliency and two-scale image decomposition. An average filter was utilized for decomposition. The extraction of visual saliency was executed using the mean and median filters. To merge the complementary information of detail images, a new weight map construction method was suggested. A comparison with seven state-of-the-art multi-scale fusion methods was implemented. The results show comparable or superior results to these methods. The implementation of the various comparable algorithms has been executed by the software provided by their authors. The comparison results of the proposed algorithm and the comparable methods are listed in Table 1. Results with the highest values are written in bolded form.

Table 1: Evaluation criteria for the test image

Clock	***MI***	$Q^{XY/F}$	***VIFF***	***SF***
Proposed	9.073	0.8935	0.943	**16.3391**
saliency	7.3824	0.8890	0.9097	13.9161
WTCV	7.0369	0.8875	0.8436	13.5187
MGFF	6.6497	0.8919	**0.9455**	13.9179
DCT	**9.1333**	**0.9015**	0.9391	14.3277
DWTPCav	7.9259	0.8566	0.8584	12.8612
Leaf				
Proposed	**7.6397**	**0.8748**	0.8113	20.4648
saliency	4.8862	0.8096	0.8600	17.3888
WT Cv	4.4872	0.8216	0.7316	15.4146
MGFF	4.1505	0.8552	**0.9738**	17.4061
DCT	7.5424	0.8697	0.7969	**20.5043**
DWPCav	5.3069	0.7721	0.7052	17.5240
Pepsi				
proposed	**9.048**	0.8958	**0.9491**	**15.5588**
saliency	7.5351	0.8545	0.8906	12.3709
Wt CV	7.1871	**0.9082**	0.8465	14.0388
MGFF	6.8561	0.8942	0.9408	12.9023
DCT	8.9450	0.8882	0.9139	12.9522
DWTpc Av	8.001	0.8918	0.8592	13.0394
Leopard				
proposed	**11.123**	0.9494	**0.9991**	**22.2942**
saliency	9.3324	0.9614	0.9437	19.1503
Wt cv	9.0714	0.9676	0.9279	19.5664
MGFF	9.0433	**0.9687**	0.9608	19.2754
DCT	11.022	0.9494	0.9975	22.2769
DWpcAV	9.0432	0.969	0.9071	19.4281
Bottle				
proposed	**8.8259**	0.8678	0.8852	**33.6681**
saliency	6.6348	0.8421	0.8667	28.7697
Wt CV	5.7381	0.8486	0.8092	28.3858
MGFF	5.9217	**0.8773**	**0.9268**	29.1702
DCT	8.7769	0.8680	0.8832	33.6646
DwPcAv	6.2779	0.7560	0.7489	26.5587
Corner				
proposed	**8.8491**	0.8764	0.9302	**20.783**
saliency	7.1549	0.8758	0.9012	17.3979
Wt CV	6.6975	**0.9148**	0.8702	17.4641
MGFF	5.8555	0.9027	**0.9573**	17.535
DCT	8.6875	0.8782	0.9194	20.5223
DWPCav	6.9924	0.8898	0.8706	17.9702

The obtained results from Table 1 show that the proposed method outperforms other comparable ones in terms of MI index except for clock images, where the DCT method is higher (0.0603). The average improvement is as follows: for *leaf* images, it is 2.36506; for *clock* images, it is 1.1471 except for the DCT method; for *Pepsi* images, it is 1.34314; for *leopard* images, it is 1.62054; and finally, for *corner* images, it is 1.77154. This means that the proposed method presents better performance in correlating between fused images and input ones. With respect to $Q^{XY/F}$, which indicates the amount of information transferred from source images to fused ones, the proposed method is similar to other comparable results. For the VIFF metric, which calculates the amount of information in the image, the proposed algorithm is the best in *Pepsi* and leopard images, and it is the second-best result in *clock, bottle*, and *corner* images. In the *leaf* test image, it is in the third step. In the case of the spatial frequency metric, which indicates the overall level of information, the proposed algorithm is preferable to other methods unless a *leaf* image is used, which lies in the second step. The average refinement is as follows: for *clock* images, it is 2.63078; for *leaf* images, it is 3.531425 except for the DCT method; for *Pepsi* images, it is 2.49808; for *leopard* images, it is 2.35478; for *bottle* images, it is 4.3583; and finally, for *corner* images, it is 2.6051. One can notice that the proposed algorithm surpasses or resembles the other methods. Also, the performance of the proposed algorithm was compared with the comparable algorithms in terms of PSNR and SSIM, as stated in Tables 2 and 3, respectively.

Table 2: PSNR (dB) of the proposed algorithm and other comparable ones

Test Images	Numbers	Clock	House	Pepper
Proposed	**43.87**	**39.05**	**40.83**	**42.09**
saliency	34.39	37.32	37.47	37.41
WtCV	33.78	35.47	36.34	35.89
MGFF	40.63	35.38	37.77	37.41
DCT	39.24	38.79	40.49	41.56
DWPCAV	35.85	35.71	35.59	35.42

Table 3: SSIM of the proposed algorithm and other comparable ones

Test Images	Numbers	Clock	House	Pepper
Proposed	**0.9657**	**0.9864**	**0.9900**	**0.9695**
saliency	0.9220	0.9186	0.8792	0.9433
WtCV	0.8831	0.8475	0.8810	0.9277
MGFF	0.8301	0.8707	0.8984	0.9492
DCT	0.9058	0.8861	0.9153	0.9671
DWPCAV	0.6986	0.8653	0.8432	0.9286

In Table (2), where a reference image is available, the results of PSNR show that the proposed algorithm offers better results than other fusion algorithms. The average improvement of the test images is as follows: for *numbers*, the image is 7.092 dB; for *clocks*, 2.516 dB; for *houses*, 3.298 dB; and finally, for *peppers*, 4.552 dB. In terms of the SSIM index, the results also reveal the superiority of the proposed method among other comparable ones. The average improvement is as follows: 0.11778 for the *numbers* image, 0.10876 for the *clock* image, 0.10658 for the *house* image, and finally 0.02632 for the *pepper* image. Since the proposed algorithm is a block fusing method that divides the input images into fixed-size blocks and chooses the appropriate one depending on activity measurement, which is the average gradient, and a consistency verification (CV) process is adopted to refine the chosen

block in order to eliminate the blocking effect, the fused images obtained from the proposed algorithm are of the best quality, have all the details of the input images, and are free of blurring and ringing effects.

6. Conclusions:

A block-based multi-focus image fusion method is proposed. This algorithm is based on applying a biorthogonal wavelet transform to the input images, equalizing the low-low subband to zero, and then applying an inverse biorthogonal wavelet transform. So the resultant images contain only edges and details. The average gradient (AG) metric is relied on for focus measurement. Each image is divided into 4 x 4 blocks, and an average gradient is calculated for every block to merge the corresponding blocks of the input images to get the focused one. Consistency verification (CV) is utilized to avoid inappropriate block choice and blocking effects in the resultant output image. Experiments were conducted on a set of multi-focus images (reference and non-reference). The resultant fused images preserve all the details and boundaries of the input images. A comparison with five state-of-the-art methods has been conducted. Four metrics for evaluation performance were used for non-reference images: MI, $Q^{XY/F}$, VIFF, and SF. The average improvements in MI and SF are 1.65 and 2.67, respectively. The proposed algorithm is similar in performance to the other methods in $Q^{XY/F}$, and VIFF. For reference images (PSNR and SSIM), metrics have been carried out. Average improvements are 4.3645 dB and 0.08986, respectively. So the proposed algorithm is effective on both objective criteria and appearance.

References

[1] S. Bhat and D. Koundal, "Multi-focus image fusion techniques: a survey," *Artif Intell Rev*, vol. 54, pp. 5735–5787, 2021.

[2] F. Citil, R. Kurban, A. Durmus, A. and E. Karaköse, " Fusion of Multi-Focus Images using Jellyfish Search Optimizer," *European Journal of Science and Technology*, vol. 37, pp. 147-155, June- 2022.

[3] O. Prakash, R. Srivastava and A. Khare, " Biorthogonal Wavelet Transform Based Image Fusion Using Absolute Maximum Fusion Rule," *IEEE International Conference on Information and Communication Technologies (ICT 2013)*, pp. 758-763, 2013.

[4] H. Li, B.S. Manjunath and S.K. Mitra, " Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing*, vol. 57, no. 3, pp. 235- 245, 1995.

[5] H. K. Abbas, A. H. Al-Saleh, H. J. Mohamad and A. A. Al-Zuky, "New algorithms to Enhanced Fused Images from Auto-Focus Images," *Baghdad Sci. J*, vol. 18, no. 1, pp. 124-131, March-2021.

[6] N. J. Habeeb, "Image Focus Enhancement Using Focusing Filter and DT-CWT Based Image Fusion," *Iraqi Journal of Science*, vol. 62, no. 9, pp. 3228-3236, 2021

[7] M. Abhyankar, A. Khaparde and V. Deshmukh, "Spatial domain decision based image fusion using superimposition," in *proceedings of the 15th IEEE/ACIS International Conference on Computer and Information Science, (ICIS 2016)*, Okayama, Japan, 2016.

[8] X. Qiu, M. Li, L. Zhang and X. Yuan, "Guided filter-based multi-focus image fusion through focus region detection," *Signal Processing: Image Communication*, vol. 72, pp. 35-46, 2019.

[9] C.R. Mohan, K. Chouhan, R.K. Rout, K.S. Sahoo, N.Z. Jhanjhi, A. O. Ibrahim and A. Abdelmaboud, "Improved Procedure for Multi-Focus Images Using Image Fusion with qshiftN DTCWT and MPCA in Laplacian Pyramid Domain," *Appl. Sci.*, vol. 12, no. 19, pp. 9495-9523, 2022.

[10] Y. Yang , S. Huang , J. Gao , and Z. Qian, " Multi-focus Image Fusion Using an Effective Discrete Wavelet Transform Based Algorithm," *Measurement Science Review*, vol. 14, no. 2, pp. 102-108, 2014.

[11] R. Srivastava and A. Khare, "Multifocus noisy image fusion using contourlet transform," *The Imaging Science Journal,* vol.6, no. 7, pp. 408-422, 2015.

[12] Y. Li, "Multi-focus Image Fusion Using Curvelet Transform," in *Proceedings of the 2nd International Symposium on Computer, Communication, Control and Automation (ISCCCA-13)*, pp. 499-504, Paris, France, 2013.

[13] Vakaimalar E, Mala K, Babu RS. "Multifocus image fusion scheme based on discrete cosine transform and spatial frequency", *Multimed Tools Appl,* vol. 78, no. 13, pp. 17573–17587, 2019.

[14] Bouzos O, Andreadis I, Mitianoudis N., "Conditional random field model for robust multi-focus image fusion", IEEE Trans Image Process, vol. 28, no. 11, pp. 5636–5648, 2019.

[15] Wang X., Hua Z., Li J., "Multi-focus image fusion framework based on transformer and feedback mechanism," Ain Shams Engineering Journal, vol. 14, no. 5, p. 101978, 2023.

[16] G.B. Gebremeskel, "A critical analysis of the multi-focus image fusion using discrete wavelet transform and computer vision," *Soft Comput* , vol. 26, pp. 5209–5225, 2022.

[17] M. Haribabu, C. Hima Bindu and K. Satya Prasad, "Image fusion with Biorthogonal Wavelet Transform based on maximum selection and region energy," in *proceeding of 2014 International Conference on Computer Communication and Informatics*, Coimbatore, India, 2014.

[18] S. Wang, T. Zhan, Y. Chen, Y. Zhang , M. Yang, H. Lu, H. Wang, B. Liu and P. Phillips, "Multiple Sclerosis Detection Based on Biorthogonal Wavelet Transform, RBF Kernel Principal Component Analysis, and Logistic Regression," *Access IEEE*, vol. 4, pp. 7567 – 7576, 2016.

[19] W. Zhou, F. Wang, X. Wang, F. Tang and J. Li, "Evaluation of Multi-Source High-Resolution Remote Sensing Image Fusion in Aquaculture Areas," *Appl. Sci*., vol. 12, no. 3, pp. 1170-1188, 2022.

[20] K. Padmavathi, C.S. Asha and M. V. Karki, "A novel medical image fusion by combining TV-L1 decomposed textures based on adaptive weighting scheme," *Engineering Science and Technology, an International Journal,* vol. 23, pp. 225-239, 2020.

[21] V. Perovic and C. Xydeas, " Objective image fusion performance characterization," in *proceedings of Tenth IEEE International Conference on Computer Vision*, Beijing, China, 2005.

[22] Y. Han, C. Yunze, Y. Cao and X. Xu, " A New image fusion performance metric based on visual information fidelity," *Information Fusion*, vol. 14, no. 2, pp. 127–135, 2013.

[23] A.A. Al-Jaburi and A.H. AL-sudani, "Medical Ultrasound Image Quality Enhancement and Regions Segmentation," *Iraqi Journal of Science*, vol. 63, no. 10, pp. 4518-4533, 2022.

[24] N.H. Hussein and M.A. Ali, "Medical Image Compression and Encryption Using Adaptive Arithmetic Coding, Quantization Technique and RSA in DWT Domain," *Iraqi Journal of Science*, vol. 63, no. 5, pp. 2279-2296, 2022.

[25] 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, pp. 5576–5605, 2019.

[26] M. Haghighat, A. Aghagolzadeh and H. Seyedarabi, "Multi-focus image fusion for visual sensor networks in DCT domain," *Computers and Electrical Engineering*, vol.37, pp. 789–797, 2011.

[27] R. Vijayarajan and S. Muttan, "Discrete Wavelet Transform Based Principal Component Averaging Fusion for Medical Images," *International Journal of Electronics and Communications*, vol. 69, no. 6, pp. 896-902, 2015.

[28] D.P. Bavirisetti, R. Dhuli, "Two-Scale Image Fusion of Visible and Infrared Images Using Saliency Detection" Infrared Physics & Technology, vol. 76, pp. 52-64, 2016.