International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 385-399, November-2023

https://doi.org/10.47001/IRJIET/2023.711052

Exploring Types of Multi-Focus Image Fusion

[1]Karam Mohammed Ghazal, [2]Dr. Ielaf O. Abdul Majjed Dahl

[1,2]Computer Science Department, University of Mosul, Mosul, Iraq

Authors E-mail: [1]karam.22csp42@student.uomosul.edu.iq, [2]ie_osamah@uomosul.edu.iq

***Abstract* - This paper explains methods. Fusion of multiple-focus images can actually take care of the profundity of field issue in optical focal point regions, the blurred picture appears strange due to the high frequency degradation Information. Most often, the camera is to blame for this. The absence of a deep field is caused by optics in the cameras. The picture becomes sharper as a result. Only in particular locations for a comprehensive focal length image, Fusion of multiple-focus images primary objective is to solve a problem with depth of field cameras. By blending at least two to some degree centred pictures into a solitary totally centred picture, Combination of various centre pictures can tackle the optical focal point's profundity of field issue.**

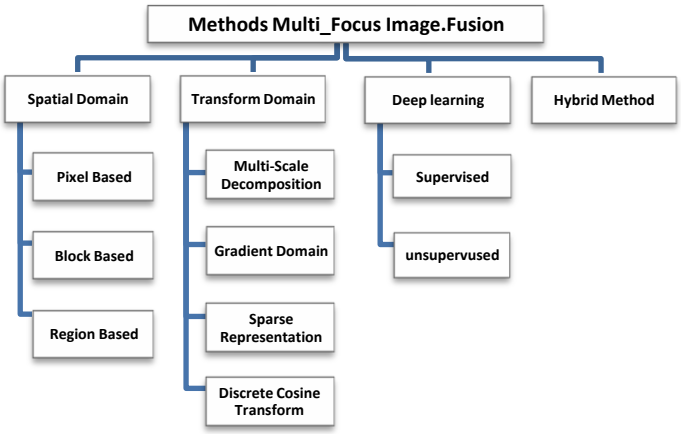
Keywords: multi-focus image fusion, image fusion, deep learning, image processing.

I. INTRODUCTION

Image fusion creates an image that is superior to the original image use a specialized application based on analysing the attributes of several photos taken simultaneously using redundant and complementary image data. Utilization of specialized methods to extract meaningful feature data from two or more photos with the use of fusion technology, a new image with more detailed and precise information can be produced image fusion can be separated into two categories based on the sorts of input source images. Remote control multi-focus picture fusion, sensor image fusion, medical image fusion, and multiple exposure image fusion combining visible and infrared images with the rise of more and more research techniques and applications. Fusion of multiple-focus images technology, for example, has a very wide range of potential applications in digital photography, computer vision.

II. LITERATURE REVIEW

Classifying image fusion techniques These methods are separated into four groups based on the currently used MFIF methods: hybrid, neural network", spatial domain, and transformation domain. The MFIF classification is displayed in Fig. 1.



1. Spatial Domain

Spatial domain fusion techniques operate directly on pixel intensity. A multifocal image's spatial characteristics are used to perform fusion. Since there is no sub-sampling process, it is also known as the single-scale fusion method. Spatial domain fusion methods fall into two categories: intensity transformation and spatial filtering. Intensity transformation describes the treatment of a single image pixel. Examples of intensity transformation techniques include image contrast manipulation and thresholding. On the other hand, spatial filtering also uses adjacent pixels in addition to individual image pixels. The two are image sharpening and image smoothing, respectively.

1.1 Pixel Based

Pixel-based techniques are the most well-known for Multi-Cantered Picture Separating (MFIF) because of their capacity to create exact pixel-wise weight maps for intertwining pictures. Nonetheless, these strategies have high aversion to commotion and misregistration, and require a bigger number of spatial neighbours to further develop dependability.

In 2003, Li et al [1]. fostered a strategy in view of pixel perceivability that works out perceivability esteem and performs combination in light of it, yet not brain organisations or fluffy rationale.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 385-399, November-2023

https://doi.org/10.47001/IRJIET/2023.711052

(In 2012, Ludusan et al)[2]. distributed a strategy for IF and denoising in light of fractional differential conditions and blunder assessment hypothesis.

In 2014, Hua et al [3]. distributed a strategy for combining different spotlight pictures in view of irregular strolls on charts. This strategy evaluated highlights like centre measures and variety consistency to make a completely associated chart. The weighting factor for DOF input pictures was determined utilising this strategy, which required less calculation, was more dependable and stable, and conquered the downsides of pixel-based procedures.

Liu et al..2015 [4]. worked on the nature of source pictures by matching misregistration pixels between numerous pictures utilizing the MFIF calculation. This strategy further developed combination characteristics for object movement and edges, yet had higher time prerequisites, memory use, and lower calculation proficiency contrasted with past methods. Itwas not material to multi-openness IF and remote detecting IF.

BAI et al. 2016 [5]. fostered a spatial technique utilising numerical morphology and slope-based choice guides, outflanking eight cutting-edge strategies quantitatively and subjectively.

Chen et al . (2007) [6] fostered an all-in-center picture strategy utilizing estimation and slope energy center, safeguarding sharpness subtleties, characterizing frontal area and foundation limits, and being strong to commotion and misregistration.

Xia et al . (2018) [7]presented a strategy for characterizing misidentified pixels into bunched and scattered bunches utilizing likelihood separating and locale revision, bringing about proficient rectification of these misidentified pixels.

Farid et al. 2019 [8]. cultivated a MFIF system using Content Flexible Clouding (Taxi), which further developed picture significance and defeated other undeniable level techniques.

Mama et al. 2019a [9]. presented an irregular strolls based technique for intertwining multi-cantered pictures, upgrading picture separation, clamour evacuation, and running time decrease.

Ji et al. 2020 .[10] developed a method using a multinomial logistic regression classifier and Random Walks to target source image objects, reducing misregistration but with low computational efficiency.

To make exact pixel-by-pixel weight maps for intertwining pictures, Multi-Center Picture Separating (MFIF) has been executed utilising pixel-based methods. To look at the lucidity of pixels in multi-centered pictures, these methods, likewise referred to as choice guides, depend on centre measures. Borders are intertwined to diminish the effect of incorrect limit accuracy, and a multi-scale mat centering strategy is proposed to incorporate centred regions Fig. 2.

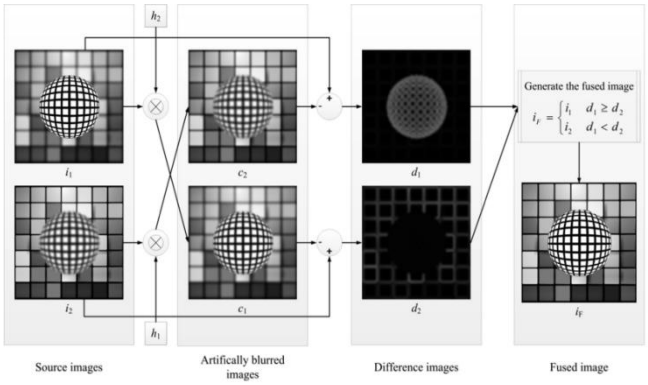


Figure 2: Fusion scheme of proposed method

1.2 Block based

The calculations for picture combinations that depend on blocks break down the source pictures into blocks of a similar size. Then, at that point, the engaged blocks are recognised as those blocks with higher centre measures from each set of blocks. Be that as it may, the adequacy of these calculations is restricted by the size of the blocks. Also, the centre measure may not necessarily, in every case, precisely recognise the completely engaged block from each pair, which can bring about antiques in the combination pictures.

Huang et al. [11] evaluated the performance of block-based fusion algorithms and found that by selecting an appropriate block size and an effective focus measure, these algorithms can produce high-quality fusion images.

Several algorithms have been proposed to address the problem of block-size selection. Aslantas et al. [12] utilized an optimization method to choose the block size, but the iterative procedures for optimization proved to be time-consuming. Additionally, alternative region-based image fusion algorithms[13] have been introduced, which involve splitting the source images into regions rather than blocks. These region-based algorithms begin by segmenting the source images using techniques such as normalized cuts[14], and then proceed to perform image fusion by measuring the clarity of corresponding regions and combining the sharply focused regions. However, the segmentation procedure often hampers the efficiency of the region-based algorithms, and the accuracy of the segmentation greatly impacts the final quality of the fusion images.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 385-399, November-2023

https://doi.org/10.47001/IRJIET/2023.711052

Li et al. (2001).[15] used spatial frequency to fuse images, which proved effective in real-time applications and outperformed wavelet transform-based methods in both qualitative and quantitative evaluation. However, the method requires further exploration for adaptive threshold selection and block size determination.

Aslantas et al. (2009)[16] developed an IF technique using Frequency Selective Weighted Median filter (FSWM) for fusing images with impulsive noise, resulting in improved image quality. The main limitation was the lack of an optimization tool.

Agrawal et al. (2010) [17]proposed a modified PCNN method for fusing multi-focused images, reducing computing time and using Energy of Laplacian (EOL) and Spatial Frequency (SF) as clarity measures.

Bai et al. (2015) [18] introduced a spatial domain MFIF method based on a quadtree structure, effectively detecting focused regions in the input images.

Vakaimalar et al. (2019) [19] investigated an IF technique combining DCT and Spatial frequency (CDSIF), achieving high fusion accuracy without blocking artifact or blurring effect.

Banharnsakun et al. (2019) [20] presented an IF method based on an artificial bee colony (ABC) algorithm, outperforming existing state-of-the-art methods by selecting sharper image blocks.

De and B. Chanda [14] The block-based approach for MFIF involves a crucial number of blocks, with small blocks containing large portions from both focused and defocused regions, while large blocks may contain small blocks and be affected by mis-registration problems Fig.3.

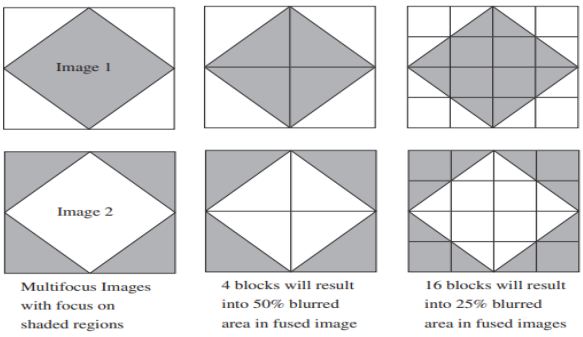


Fig. 2. Multi-focus image fusion with equal-sized blocks.

1.3 Region based

Scientists have contrived district-based strategies to accomplish an adaptable division of the source picture. Both the district-based and block-based spatial techniques share a comparative system. Notwithstanding, the key differentiation lies in the way that district-based strategies assess the movement level in sporadically measured, divided areas. Tragically, the area-based technique is tormented by impeding impacts, which is its disadvantage.

One way to deal with multi-center picture combination includes utilizing a locale place-based part, which is a numerical capability that decides the significance of every pixel in the information pictures for the last melded picture. This part allocates higher loads to pixels that are nearer to the focal point of the locale of interest, while giving lower loads to pixels that are farther away. By changing the piece, it is feasible to control the size, shape, and perfection of the progress between the engaged and defocused areas.[21]

To start with, various photographs of a similar scene at different centre distances should be taken to execute a multi-focus image combination utilising a local focus-based portion. To gauge every pixel's commitment, the portion would then be applied to each picture. The next stage is to meld the weighted pictures together, utilising a suitable combination technique, for example, averaging or weighted averaging, to make the intertwined image.

The district-based piece is a scaled projection of the middle pixel. The clarification of the area-focused portion and the naming method Figure 6 shows the underlying centre guide.

The creation of the initial focus map is shown in the block diagram below. It is created using a region-based, central kernel. A geographical region or all-one matrix with a size of 3*3 is produced when the sliding window reads the neighborhood's information. When compared to typical sliding windows, which merely output a pixel value as the center pixel, it is extremely different. Additionally, it can manage noise sensitivity and unstable pixels thanks to the region-center based kernel. Given that the task is straightforward, quick, and precise, this procedure is essential to the algorithm.

High stability and accuracy may be seen in the resulting focus map. The algorithm's subsequent steps will be straightforward and inexpensive.[22]

The sliding window's pixel power is utilised by the channel to think about the surfaces of all info pictures in the wake of figuring out and contrasting the pixel force of all info pictures. The assessed centre slope will result from this correlation. To distinguish the central locale, this guide turns to the critical data that should be examined.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 385-399, November-2023

https://doi.org/10.47001/IRJIET/2023.711052

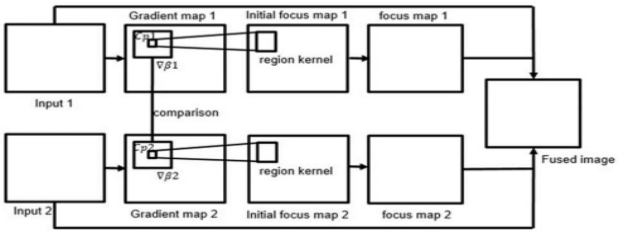


Figure 4: The block diagram of the method

The region-center based kernel is more effective. It operates more expensively than pixel kernel since it is a regional scale. The longer processing times on a computer, however, are not greatly affected by the higher cost of processing. Computers today typically operate at high technical levels. [23]

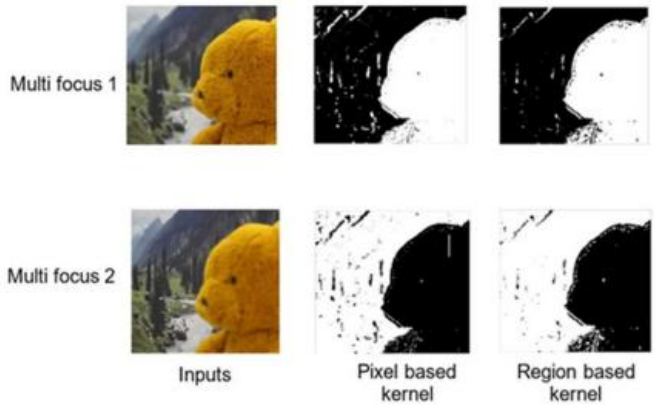


Figure 5: Initial focus region map from multi-focus images

In Fig. 7, the initial focus region according to the region-based and pixel-based kernels is compared. The input pictures feature a variety of bear toys in the background textures. Both the in-focus and out-of-focus areas are remarkably similar. The noise is kept there by the pixel-center-based kernel. However, the region-center-based kernel has the ability to transform background noise into minute objects.

II. TRANSFORM DOMAIN

2.1 Multi-Scale Decomposition

The advantage of "multi-scale decomposition" is that it allows the analysis of data at multiple levels, capturing both global trends and local details. This approach is particularly useful in applications such as image processing, where images often contain structures at different scales, such as edges, textures, and objects of different sizes.

By decomposing data into different scales, it becomes possible to analyze and manipulate specific components independently. This can facilitate tasks such as denoising, compression, feature extraction, and pattern recognition.

Overall, multi-scale decomposition provides a powerful framework for understanding and processing complex data by revealing its underlying structures at different scales or levels of detail.

- **Visual saliency detection**

Saliency detection (SD) [24] is the technique of visually identifying or differentiating prominent regions , such as pixels or objects , that attract more visual attention from viewers compared to other parts of the image.

- "Multi-focus image" datasets

- Saliency maps of datasets of multiple-focus images

A few of the approaches discussed in [25] and [26] produce saliency maps with a poor level of resolution. The bounds of certain additional SD methods in [27] and [28] are not well defined. Producing saliency maps for fusion purposes is rendered useless due to these restrictions, so Achanta et al., introduced the frequency-tuned saliency detection technique (FTSD) to overcome the limitations of existing SD approaches. This algorithm is capable of meeting every need for an effective SD approach. However, if an image has a complicated backdrop or huge, conspicuous items, the algorithm fails. Achanta et al. developed the maximum symmetric surround SD approach (MSSS) to address these problems.

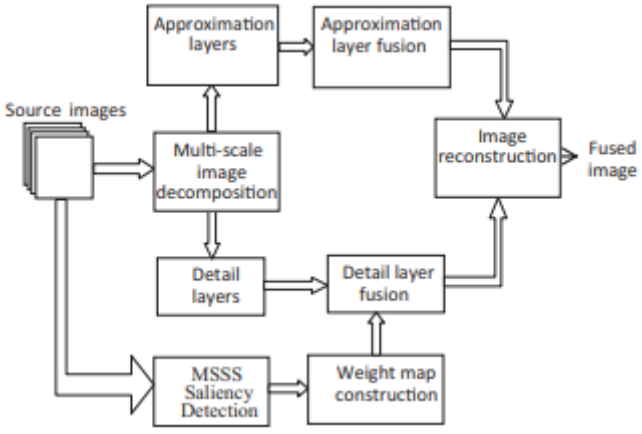


Figure 6: General block diagram of the proposed algorithm

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 385-399, November-2023

https://doi.org/10.47001/IRJIET/2023.711052

SD algorithms are used. We note that salient portions of the multi-focus images can be extracted using the MSSS saliency detection algorithm [29]. The following reasons are why we favor MSSS above other SD methods:

(1) It generates saliency maps of high resolution that have clearly defined borders.

(2) It effectively emphasizes the important parts of images that have a complicated background.

The MSSS detection algorithm theory involves deriving a saliency map by calculating the Euclidean distance between the mean of an image I_μ and each pixel of the "Gaussian blurred image,$I_f(u,v)$producing the MSSS saliency map" for a "particular sized" image.

The (MSSS) saliency map is determined for an image I with dimensions w and h by the following method.

$$S_{ss}(u,v) = \|I_\mu(u,v) - I_f(u,v)\|,$$

Where $I_\mu(u,v)$ the subimage mean, centered at pixel (u,v) can be represented as follows:

$$I_\mu(u,v) = \frac{1}{A}\sum_{i=u-u_0}^{u+u_0}\sum_{j=v-v_0}^{v+v_0} I(i,j)$$

whereu_0; v_0A denotes the area calculated as follows, indicating off-sets.

$$u_0 = \min(u, w-u)$$

$$v_0 = \min(v, h-v)$$

$$A = (2u_0+1)(2v_0+1)$$

The sub-images from earlier mathematical operations are the maximum symmetric surround regions for a given central pixel, while the focused regions in multi-focus photographs contain more visual information compared to the defocused regions. We can use SD algorithms to identify the prominent areas in the out-of-focus photos and extract the salient areas of the multi-focus pictures using the MSSS saliency detection technique. The process of extracting visual saliency in MSSS is denoted as follows:

$$S = \text{MSSS}(I)$$

The source image is denoted as I, while its saliency map is denoted as S.

The N-images are decomposed into layers with some layers capturing small scale intensity variations and others capturing large scale intensity variations:

$$B_n^{k+1} = B_n^k * A, \text{where }, k = 0,1,\dots,K$$

The present level K+1 is achieved by subtracting the approximation layers $B_n^k$from the previous level k from the approximation layers B_n^{k+1}at the present level K+1.

$$B_n^{k+1} = B_n^{k+1} - B_n^k .$$

Visual saliency detection The MSSS detection algorithm is used to obtain the visual saliencies of multi-focus images , and the process of extracting saliency from the approximation layersB_n^k at k levels is represented as follows:

$$S_n^{k+1} = \text{MSSS} \mid (B_n^k)$$

The saliency map of the nth source image, referred to as Sn, was validated using visual saliencies from a flower dataset , with (k = 3).

Weight map calculation" The focus region for each source image in MFF is specified in great detail, and a single image can be generated by merging the desired areas from all the photographs. To achieve this, appropriate weight maps can be applied to the input images, which can distinguish between focused and defocused parts. These weight maps are obtained by normalizing the saliency maps.

$$w_i^{k+1} = \frac{S_i^{k+1}}{\sum_{n=1}^{N} S_n^{k+1}}, \forall i = 1,2,\dots,N$$

The recommended weight maps for a flower dataset can recognize centered and defocused districts, with the red and green square shapes showing the engaged and obscured regions in the source photographs, and these weight maps are intended to be correlative.

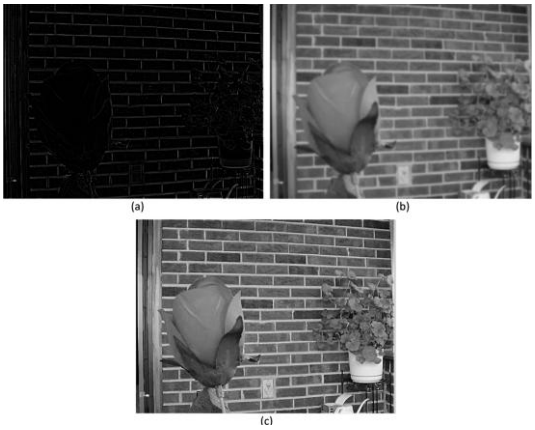
The last detail layer D^- (Fig. 7(a)) is inferred through the course of the detail layer fusion.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 385-399, November-2023

https://doi.org/10.47001/IRJIET/2023.711052



(Fig. 7) The last estimation, detail layers, and combined picture are displayed in a visual presentation: (a) the last detail layer, (b) the last estimate layer, and (c) the proposed SDMF melded picture.

$$\bar{D} = \sum_{k=1}^{K} \sum_{n=1}^{N} w_n^k D_n^k$$

The final detail layer is obtained by assigning weight w_n^k to the detail layers D_n^k in a clear manner.

- The fusion of the approximation layer

The final approximation layer (Figure 7 (b)) is obtained by taking the average of the approximation layers in the following manner:

$$\bar{B} = \frac{1}{NK} \sum_{k=1}^{K} \sum_{n=1}^{N} B_{n*}^k$$

Reconstruction of a fused image.

The fused image F (Fig.7 (c)) is created by merging the final base (B) and final detail (D) layers in a specific manner. [30]

$$F = \bar{B} + \bar{D}$$

2.2 Gradient Domain

The GD-based method combines the gradient representation of the source image while limiting the gradient of the fused image to a specific threshold requiring the gradient information of the image component. To ensure smoothness in the gradient domain, the Poisson equation was solved at each resolution, as demonstrated by Paul et al. s work [31].

Wang et al [32] proposed a technique for image fusion using the structure tensor where the source images are stacked into a multi-valued image and the structure tensor of each source image is computed based on its gradient graph.

The visual impact of an image can be improved and the specifics and structural data of the source image can be preserved by using the technique of gradient domain image fusion, which allows for both multi-focus image fusion and multi-exposure image fusion.

Many other transform-based multi-focus image fusion techniques like independent component analysis (ICA), compressed sensing (CS), high order singular value decomposition(HOSVD), discrete cosine transform (DCT), and cartoon-texture decomposition (CTD) have also been effectively used.

Hong et al. [33] presented a method in which the preservation of salience in GD served as the basis. This technique involves creating a saliency map for the input images and correctly focusing on gradients with higher saliency values in the target gradient. The applicability of this approach to colored images was also explored. However, the enhancement of local images was not accomplished using this particular technique. Additionally, Piella et al. [34] introduced a GD-based MFIF technique that utilized a structure tensor to describe the input image geometry (Piella 2009).

This technique selects a weighted structure tensor as its method. By employing this approach, it becomes possible to eliminate blurring, ringing, and haloing artifacts. Subsequently, Sun et al. examined an MRF GD-based approach for MIFIF (Sun et al. 2013) [28]. Utilizing the Poisson equation is one of the applications of reassembling the combined image. The outcomes achieved from this technique surpassed those obtained from competing methods.

(Zhou, Yang. 2021) [35]A gradient-based approach is introduced to generate an all-in-focus image using a convolution neural network (CNN). The method inputs original images and gradient images into five models, generates initial focus score maps, and merges them to form a fused image. The dictionary constructing algorithm is a crucial tool in sparse coding, determining signal representation ability. Two methods are available for offline access: using analytical models like over-complete wavelets and curvelets, or applying machine learning techniques from numerous training image patches. The former is simple but not adaptive for complex image structures.

2.3 Sparse Representation

The problem of image fusion relying on local information of source images is addressed by the sparse-representation based multi-focus image fusion approach. It achieves the

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 385-399, November-2023

https://doi.org/10.47001/IRJIET/2023.711052

sparse representation shift invariant through a sliding window approach and divides the source pictures into small segments using a fixed dictionary of limited size. The algorithm is divided into three sections: creating the dictionary, representing the image, integrating, and reconstructing.[36]

Adaptively extracting source image patches from source image patches is done using the K-SVD [37]algorithm. To preserve each source image signal, a joint dictionary is built, and the batch-OMP algorithm is used to estimate the coefficient vectors. Utilizing a maximum weighted multi-norm-based fusion rule, fused coefficients are produced. Rebuilding the output image with fused coefficients and combined dictionary in Fig. 8(b).

- Dictionary constructing

The dictionary constructing algorithm is an important tool for assessing the signal representation capacity of sparse coding. There are two ways to access the data offline: either by applying machine learning techniques from multiple training image patches, or by using analytical models such as over-complete wavelets and curvelets. The former is straightforward but unsuitable for intricate image structures.[37] The sparse model depends on an over-complete word reference, which can be acquired through two techniques: pre-developing a word reference utilizing logical techniques like DCT, wavelets, and curvlets, or gaining a word reference from various picture patches utilizing a preparation calculation like MOD or K-SVD. Yang applied the inadequate portrayal hypothesis to the picture combination. Meagre portrayal is utilised in different examinations, including a clever system for concurrent picture combination and super-goal, which utilises 6,000 patches from six pictures to learn word references .Liu Y, Wang Z (2015)[38] proposes a multi-center picture combination strategy utilising a data set of forty great normal pictures, while Aharon offers a K-SVD-based procedure for preparing an excess word reference on a gathering of pictures.

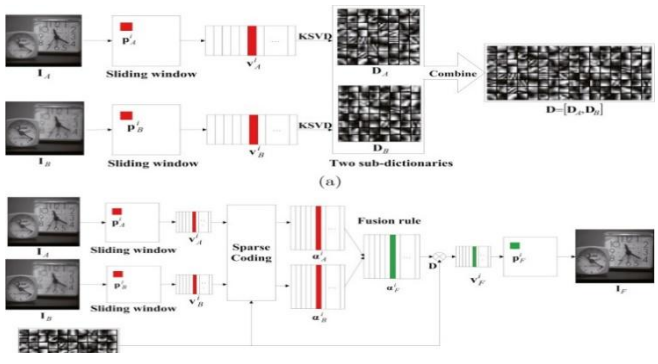


Figure 8The structure sparse-representation-based multi-center methodology (a) Strategy of the proposed word reference learning technique (b) Outline of the multi-center picture combination approach

The sliding window procedure is utilised to catch neighbourhood remarkable elements in two enrolled source pictures (IA and IB) with size M × N. This strategy partitions each source picture into patches of size n × n, which are changed into vectors by means of lexicographic requesting. These vectors structure a framework VA, where every section compares to one fix in the source picture IA.

SR strategies have arisen as a critical branch in the field of MFIF, In 2010, Yang et al. presented a SR-based MFIF technique (Yang and Li 2010).[39] The source pictures are addressed by utilising meager coefficients, and the picture is remade by applying the most extreme combination rule to these coefficients. This technique outperformed different existing strategies like DWT, NSCT, curvelet change (CVT), SWT, morphological wavelet change (MWT), and spatial slope(SG). Consequently, Yang et al. fostered a SR-based strategy for IF of multi-centered pictures (Yang and Li 2012).[40] This technique uses the synchronous symmetrical matching pursuit (SOMP) for picture combination, permitting it to deal with pictures with clamour. In any case, it ought to be noticed that this strategy is very tedious.

Chen and colleagues (2013).[41] introduced a method for MFIF that focuses on super-resolution (SR) and region-based techniques. The SR coefficients are utilized to construct an enhanced clear image, enabling the extension of the depth of field (DOF) and the generation of a fused image with enhanced clarity. However, this method did not explore structured SR.

In a subsequent study, Liu and Wang (2015).[38] presented a MFIF method based on Adaptive SR (ASR). This approach involves learning compact sub-dictionaries that aid in both image fusion and denoising. Notably, this ASR-based method outperformed traditional SR-based methods both qualitatively and quantitatively.

Yin and Partners (2016)[37] further added to the field by proposing a SR-based MFIF strategy that builds a total versatile word reference. This strategy consolidates versatile sub-word references utilizing the K-particular worth decay (K-SVD) calculation and uses a greatest weighted multinorm combination rule for picture remaking.

Additionally, Ma and colleagues (2019b).[42] reported a MFIF method that combines optimal solution and joint SR. The adaptive dictionaries obtained through K-SVD are combined with fixed dictionaries, and redundant components are removed using SR. While this method yields satisfactory fusion results, it does come with a significant computational time requirement.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 385-399, November-2023

https://doi.org/10.47001/IRJIET/2023.711052

2.4 Discrete Cosine Transform

The discrete cosine change (DCT) is a generally utilised change in picture pressure, shaping the reason for different business guidelines like H263 video coding, movement JPEG, MPEG, and actually picture coding. DCT is a change space strategy zeroing in on low-recurrence parts in source pictures, with different creators detailing research on Discrete Cosine Change for MFIF

In2006, Zafar et al.[43] led an examination concerning a combination procedure for multi-centered and multi-openness pictures utilising the DCT space (Zafar et al. 2006). This approach was likewise pertinent to pictures acquired from multi-openness and could be executed in camera for shaded pictures. Therefore, Haghighat et al. introduced a plan for MFIF in view of DCT (Haghighat et al. 2010, 2011).[44] They determined the difference in theDCT to foster a constant combination technique for MFIF, which brought about better picture quality. Besides, the intricacy of this continuous application was diminished. Be that as it may, this technique had a few shortcomings regarding limits and centred regions. Furthermore, Phamila et al. proposed a MFIF technique in light of DCT for consolidating multi-center pictures (Phamila and Amutha 2014).[45] This technique was both energy-proficient and incredibly straightforward. One significant downside of involving DCT is its intricacy and slight shortcoming in dealing with limits.

The idea of spatial recurrence, initially acquainted with the workings of the human visual framework, shows the degree of movement inside a picture overall. Notwithstanding the ongoing test of grasping the physiological strategies for the human visual framework, spatial recurrence remains a powerful model for picture combination because of its capacity to make proficient differences. The simplicity with which spatial recurrence can be determined in the DCT space is obvious, considering the estimation of the difference between blocks of source pictures utilizing this value.

When compared to other image fusion approaches in the DCT domain, such as DCT + Average[46], DCT + Contrast[44], DCT + AC – Max[47], DCT + Variance[48], and DCT + Variance + CV ,Multi-scale based fusion techniques like DWT, SIDWT , and NSCT are regarded as cutting-edge methods.

III. DEEP LEARNING

3.1 Supervised

Liu and colleagues were the pioneers in introducing a novel approach to MFIF using a Convolutional Neural Network (CNN) (Liu et al., 2017a)[49]. Their method involved creating a direct correspondence between input images and the focus map. Through training a CNN model, they were able to generate fusion rules and measure activity levels simultaneously, thereby addressing the challenges encountered by other fusion methods.

Li and colleagues (2020)[50] proposed the Deep Regression Pair Learning (DRPL) model, which is based on deep learning, for MFIF. In conventional end-to-end CNN models the input images are fragmented into small patches and information is extracted from these patches. However, in this approach, the entire image is transformed into a binary mask instead of using patches, and the fusion of these masks is performed.

Zhang et al. presented a method called IFCNN, which is an Image Fusion scheme based on Convolutional Neural Networks (CNN) (Zhang et al. 2020).[51] This approach involves extracting image features from input images using two convolutional layers. These extracted features are then fused together to obtain fused features. Finally, the fused features are reconstructed to generate the resultant image, which contains more precise information compared to the input images.

In the year 2019, Lai and colleagues published a MFIF technique that relied on a deep convolutional neural network model called MADCNN (Lai et al. 2019).[52] This approach proved to be effective in addressing the issue of accurately distinguishing between the areas of an input image that were out of focus and those that were in focus.

Naji et al.[53] introduced fusion algorithms, known as ECNN, which differ from the previously mentioned methods by employing three CNNs instead of a single model. The underlying concept is to leverage multiple models and datasets, rather than relying solely on one, in order to mitigate the issue of over fitting on the training dataset.

HCNN, another technique, was introduced by Naji et al. [54] The key innovation of ensemble learning-based approaches lies in their ability to integrate the strengths of multiple models or datasets, making them more resilient to diverse inputs.

In addition to the supervised methods mentioned above, there have been other supervised approaches suggested as well. For instance, Zhai et al.[55] Introduced an MFIF approach that relies on the denoising auto-encoder (DAE) and deep neural network (DNN). Deshmukh et al. [56] put forward an algorithm that calculates weights to identify the sharp areas in input images by utilizing the Deep Belief Network (DBN). Additionally, Lahoud et al.[57] proposed the utilization of pre-

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 385-399, November-2023

https://doi.org/10.47001/IRJIET/2023.711052

trained neural networks to extract features, thereby easing the training phase.[58]

3.2 Unsupervised

Xu and colleagues[59] presented a method called Fusion DN, which intends to consolidate different picture combination undertakings into a firm and interconnected network. To keep the model from ignoring the information gained from earlier errands during successive preparation, they executed flexible weight union. Therefore, a particular model is produced that can be used for different combination errands. Additionally, Fusion DN consolidates SSIM in the misfortune capability as well as considers the perceptual misfortune and angle misfortune.

Additionally, Xu et al. [59] proposed a choice guide put together for solo MFIF calculations based on slopes and associated locales, named GCF. Unique in relation to SESF and GCF utilizes an encoder-decoder organization to create a guide Mo and the underlying double cover M1, which will be utilized to compute the misfortune in light of slope and the misfortune in view of associated districts, separately. Moreover, GCF requires multi-center picture matches when preparing information. A post-handling step, for example, a consistency check, is then performed to get an official conclusion map (MF).

In addition to end-to-end approaches, a few unsupervised MFIF methods based on decision maps have also been suggested. An example is the unsupervised MFIF algorithm proposed by Ma et al. [60], which utilizes an encoder-decoder network (SESF) and includes SSIM as part of the loss function. During the inference phase, SESF employs the encoder to extract deep features from input images. These features, along with spatial frequency, are then used to obtain the focus map. Finally, consistency verification methods are applied to generate the ultimate decision map.

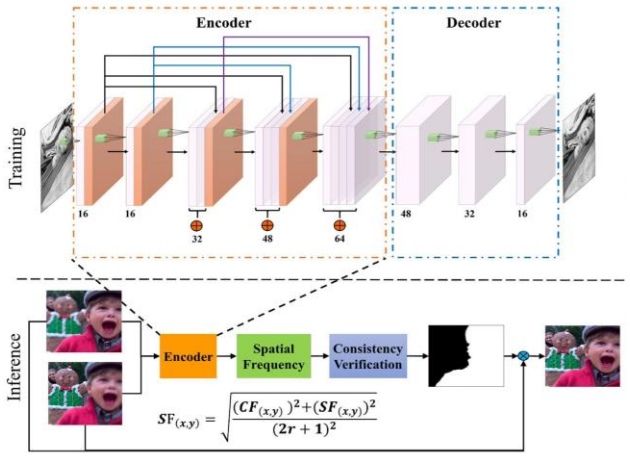


Figure 9: The structure of the SESF network architecture

GANs" have also been applied in unsupervised MFIF approaches recently. The first GAN-based" unsupervised "MFIF method" is MFF-GAN"[61]. In this strategy, a versatile choice block is first used to assess the sharpness of every pixel in the source pictures, utilizing the rehashed obscure standard. In particular, in the event that a pixel has higher sharpness, its worth changes more in the wake of adding obscure. Then, a substance misfortune is explicitly intended to guarantee that the generator creates a melded outcome with a similar dispersion as the center source pictures. At long last, a "discriminator" is utilized to make an ill-disposed game with the generator" meaning to make the slope guide of the melded picture like the joint inclination map developed from the source pictures. This further enhances the surface subtleties.

One more way to deal with picture combination, known as the PMGI, was proposed by a similar gathering and offers a brought-together answer for different picture combination undertakings. As opposed to Fusion DN, PMGI tends to the errand of picture combination by considering the surface and force upkeep issues of source pictures. The PMGI model is partitioned into two ways: the power way and the inclination way. The misfortune capability of PMGI comprises both a power part and a slope part.

The U2Fusion strategy, which was depicted prior, was developed [62] with two tremendous changes. First and foremost, rather than utilizing the source pictures for allotting the level of data conservation, the task was done in light of the estimation of data from the separated elements. Besides, the misfortune capability was adjusted all the while

$$L = \text{lssim} + \alpha \,\text{lm}\, s \,e$$

The mean squared error (MSE) between two pictures describes LMS. It was shown by them that the basic change can assist with getting the focal qualities of the source pictures, while the following change can uphold diminishing the luminance difference in the joined picture [62].

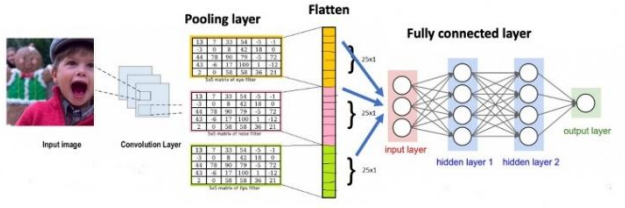


Figure 10: CNN (convolutional neural network)

IV. HYBRID METHODS

The advantages and disadvantages of the MFIF field apply to both spatial and transform domain techniques. Combining the benefits of these approaches, hybrid-based

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 385-399, November-2023
https://doi.org/10.47001/IRJIET/2023.711052

methods have been developed to create more efficient IF approaches that can tackle the limitations of traditional methods. As deep learning is one of the rapidly advancing branches of MFIF, certain spatial or transform domain approaches have been enhanced to enhance fusion outcomes through integration with deep learning techniques.

Li et al. introduced a technique in 2013 that involved a dual window methodology and the detection of focused regions using an IF method [63]. The enhancement of the Multiscale Top-Hat (MTH) transform allowed for the identification of concentrated regions within the image a dual window approach was used in combination with MST to address the gaps in the transition region. This technique effectively overcame the limitations of the spatial domain. In addition, (Li et al) [64]. showcased two MFIF techniques that utilized multi-scale and multi-direction neighbors distance. These techniques were demonstrated in their study. In order to enhance fusion performance, two update strategies were implemented.

This strategy performed better when contrasted with different techniques. Nonetheless, it had a critical disadvantage in that it didn't use remote-detecting methods like infrared and noticeable imaging or clinical imaging.

Li et al. proposed a technique called MFIF, which utilized a multi-scale neighbor method [65]. In this study, IF was implemented in three stages. Prior to extracting focus information, the image underwent an initial analysis using a multiscale neighbor approach. The outcome of this process included the generation of weight maps and decision maps.

In2018, Lia et al. introduced a method to combine noisy images with different focus points by employing Low-Rank Representation (LRR). This approach entailed manipulating the spatial frequency to merge the low-frequency elements and utilizing the LRR coefficients to fuse the high-frequency components. The resulting images were successfully merged [66].

A better combination execution was accomplished for photographs with commotion. The given technique has beated any remaining present status of craftsmanship strategies by accomplishing the most elevated upsides of RMSE, PSNR, and SSIM on different sorts of commotion, like salt-and-pepper, Gaussian, and Poisson clamor. To further improve results for both enlisted and misregistered pictures, Yang et al.[67] fostered a technique for MFIF by consolidating strong inadequate portrayal with a versatile PCNN. In spite of the previously outflanking present status of workmanship methods, the productivity of this innovation should have been moved along. He and his team proposed a technique called MFIF which is based on dividing the focus region into several parts in the NSCT domain, along with PCNN.[68] The resulting fused image from this approach contains a greater number of clearly defined pixels compared to the original image. Furthermore, it provides supplementary details and enhances the depth of information.

Hou et al. introduced an MFIF technique in their publication. [69] which focused on fusing colorful pictures. This approach effectively preserved the edges of the decision maps by utilizing the non-Sub sampled Shearlet Transform and KN earest Neighbors matting. Following this, Li et al. developed a novel MFIF method that combined spatial and transform-domain techniques [7].

V. EVALUATION METRICS

Quality assessment of MFIF calculations is a provoking undertaking because of the absence of ground truth in MFIF datasets. There are two normal ways to deal with the nature of the melded picture: abstract assessment and objective assessment. Abstract assessment, otherwise called subjective assessment, includes human onlookers outwardly surveying the presentation. This is profoundly important in MFIF research. Be that as it may, emotional judgement has its constraints. First and foremost, it can't be mechanised, making it tedious to assess each melded picture. Also, the assessment might be one-sided, as every observer has their own norms. In this way, abstract assessment is ordinarily joined with true assessment in picture combination research. Objective assessment, additionally alluded to as quantitative assessment, measures the combination execution utilizing assessment measurements. Various evaluation measurements have been developed, like cross entropy (CE) [70], spatial recurrence (SF), and standardised common data (NMI) [71]. Every measurement principally assesses MFIF calculations according to a particular viewpoint. Thus, it is pivotal to utilise different kinds of measurements to assess MFIF calculations.

There are four classifications of picture combination measurements.

- Data-hypothesis-based measurements
- Picture include based measurements
- Picture primary similitude-based measurements
- Human insight-propelled measurements

1) Data-hypothesis-based measurements

The assessment measurements given in this segment are established on normally used standards in data hypotheses, like entropy, common data, and motion towards commotion proportion. EN is only utilised in these estimations to evaluate the data included in the blended picture. The leftover

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 385-399, November-2023

https://doi.org/10.47001/IRJIET/2023.711052

measurements are applied to both source pictures and the blended picture to survey the amount of data moved from source pictures to the combined picture or to check the connection between the source pictures and the consolidated picture. These measurements assess the presentation of picture combinations in different ways by utilising assorted ideas or definitions. For example, while CE, QN ICE, and TE all depend on entropy, CE uses cross entropy, QN ICE is intended to check nonlinear connections, and TE utilises Tsallis entropy.

2) Picture include based measurements

1. The average gradient (AG): is a metric that quantifies the gradient data in the merged image, which in turn reflects its intricacy and texture [72].

2. Edge intensity (EI): quantifies the level of edge intensity present in an image [73]. A greater value of EI signifies enhanced clarity and superior image quality. Sobel operator [74] can be employed to calculate EI.

3. Edge-based similarity measurement (QAB/F): quantifies the extent to which edge information is conveyed from the source images to the fused image [75].

4. Standard deviation (SD): The distribution and contrast of the fused image can be observed through the SD [76]. The human visual system is highly responsive to contrast, which means that areas in an image with high contrast tend to capture human attention. It is worth noting that a fused picture with strong contrast will result in a larger SD, indicating a more pleasing visual impact in the fused image.

5. Spatial frequency (SF): SF has the ability to assess the distribution of gradients in an image, which in turn exposes the intricacy and texture within the image [77]. This signifies the presence of well-defined edges and textures, ultimately implying a commendable fusion performance.

3) Picture primary similitude-based measurements

1) Structural similarity index measure(SSIM).

SSIM is used to model image loss and distortion, thus reflecting the structural similarity between images [78].

2) Yans metric (QY).

QY serves as a fusion quality metric that is based on SSIM [79]. This metric measures the extent to which the fused image F preserves the structural information from the source images. A higher QY value signifies that more information from the source images is maintained in the fused image, thereby indicating superior fusion performance. The maximum value that QY can attain is (1).

4) Human insight-propelled measurements

1) The measurement of human visual perception (QCB) primarily assesses the similarity of the main characteristics in the human visual system [80]. y. A higher QCB value represents a greater preservation of information from the source images in the fused images, thereby indicating an improved fusion performance. The value falls within the range of [0,1].

2) Visual information fidelity, also known as VIF, assesses the fidelity of the fused image in terms of information, aligning it with the capabilities of the human visual system [81].

Also, as both source pictures contain fundamental information, different evaluation rules are made to measure the closeness between the interlaced picture F and the sources. A convincing picture mix system should successfully organize basic data from both source pictures into the last merged picture.

VI. CONCLUSION

MFIF makes it possible to combine several images with various focal planes to create a single, focused image. As a result, this method combines multiple multi-focused images into one higher-quality picture. Even though the literature has published a number of MFIF techniques for fusing defocused images, there are still some shortcomings in the state-of-the-art techniques that could be addressed. This study has addressed this by developing a novel classification system to classify various MFIF approaches. These strategies include transform domain, deep learning, and spatial domain methods as well as their hybrids It is important to develop a reliable algorithm or technique to merge two or more multi-focus images. In addition to producing good fusion, this method or algorithm should be adaptable enough to work with different datasets and improve the imaging schemes depth of field. The results of this study will be helpful to researchers in the future who want to create new MFIF techniques with deeper field coverage. Subsequent investigations may concentrate on developing a unique MFIF methodology that can successfully address current issues.

Conflict of Interest

It must be acknowledged that the authors involved in this study, as described in this publication, do not have any personal or scientific conflicts of interest that could potentially influence the findings.

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 385-399, November-2023
https://doi.org/10.47001/IRJIET/2023.711052

Data Availability

Data sharing is not applicable to this article as no new datasets were generated or analyzed during the current study.

REFERENCES

[1] Z. Li, Z. Jing, G. Liu, S. Sun, and H. Leung, "Pixel visibility based multifocus image fusion," in *International Conference on Neural Networks and Signal Processing, 2003. Proceedings of the 2003*, 2003, vol. 2, pp. 1050–1053.

[2] C. Ludusan and O. Lavialle, "Multifocus image fusion and denoising: a variational approach," *Pattern Recognit. Lett.*, vol. 33, no. 10, pp. 1388–1396, 2012.

[3] K.-L. Hua, H.-C. Wang, A. H. Rusdi, and S.-Y. Jiang, "A novel multi-focus image fusion algorithm based on random walks," *J. Vis. Commun. Image Represent.*, vol. 25, no. 5, pp. 951–962, 2014.

[4] Y. Liu, S. Liu, and Z. Wang, "Multi-focus image fusion with dense SIFT," *Inf. Fusion*, vol. 23, pp. 139–155, 2015, doi: https://doi.org/10.1016/j.inffus.2014.05.004.

[5] X. Bai, M. Liu, Z. Chen, P. Wang, and Y. Zhang, "Multi-focus image fusion through gradient-based decision map construction and mathematical morphology," *IEEE Access*, vol. 4, pp. 4749–4760, 2016.

[6] Y. Chen, J. Guan, and W.-K. Cham, "Robust multi-focus image fusion using edge model and multi-matting," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1526–1541, 2017.

[7] X. Xia, Y. Yao, L. Yin, S. Wu, H. Li, and Z. Yang, "Multi-focus image fusion based on probability filtering and region correction," *Signal Processing*, vol. 153, pp. 71–82, 2018.

[8] M. S. Farid, A. Mahmood, and S. A. Al-Maadeed, "Multi-focus image fusion using content adaptive blurring," *Inf. fusion*, vol. 45, pp. 96–112, 2019.

[9] J. Ma, Z. Zhou, B. Wang, L. Miao, and H. Zong, "Multi-focus image fusion using boosted random walks-based algorithm with two-scale focus maps," *Neurocomputing*, vol. 335, pp. 9–20, 2019.

[10] Z. Ji, X. Kang, K. Zhang, P. Duan, and Q. Hao, "A two-stage multi-focus image fusion framework robust to image mis-registration," *IEEE Access*, vol. 7, pp. 123231–123243, 2019.

[11] Y. Huang, W. Li, M. Gao, and Z. Liu, "Algebraic multi-grid based multi-focus image fusion using watershed algorithm," *IEEE Access*, vol. 6, pp. 47082–47091, 2018.

[12] V. Aslantas and R. Kurban, "Fusion of multi-focus images using differential evolution algorithm," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 8861–8870, 2010.

[13] L. Zhang, G. Zeng, and J. Wei, "Adaptive region-segmentation multi-focus image fusion based on differential evolution," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 33, no. 03, p. 1954010, 2019.

[14] I. De and B. Chanda, "Multi-focus image fusion using a morphology-based focus measure in a quad-tree structure," *Inf. Fusion*, vol. 14, no. 2, pp. 136–146, 2013, doi: https://doi.org/10.1016/j.inffus.2012.01.007.

[15] S. Li, J. T. Kwok, and Y. Wang, "Combination of images with diverse focuses using the spatial frequency," *Inf. fusion*, vol. 2, no. 3, pp. 169–176, 2001.

[16] V. Aslantas and R. Kurban, "A comparison of criterion functions for fusion of multi-focus noisy images," *Opt. Commun.*, vol. 282, no. 16, pp. 3231–3242, 2009.

[17] D. Agrawal and J. Singhai, "Multifocus image fusion using modified pulse coupled neural network for improved image quality," *IET image Process.*, vol. 4, no. 6, pp. 443–451, 2010.

[18] X. Bai, Y. Zhang, F. Zhou, and B. Xue, "Quadtree-based multi-focus image fusion using a weighted focus-measure," *Inf. Fusion*, vol. 22, pp. 105–118, 2015.

[19] E. Vakaimalar and K. Mala, "Multifocus image fusion scheme based on discrete cosine transform and spatial frequency," *Multimed. Tools Appl.*, vol. 78, pp. 17573–17587, 2019.

[20] A. Banharnsakun, "Multi-focus image fusion using best-so-far abc strategies," *Neural Comput. Appl.*, vol. 31, pp. 2025–2040, 2019.

[21] Q. Li, J. Du, F. Song, C. Wang, H. Liu, and C. Lu, "Region-based multi-focus image fusion using the local spatial frequency," in *2013 25th Chinese control and decision conference (CCDC)*, 2013, pp. 3792–3796.

[22] K. Hawari, "Multi Focus Image Fusion with Region-Center Based Kernel," vol. 11, no. 1, 2021.

[23] S. Li and B. Yang, "Region-based multi-focus image fusion".

[24] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *2009 IEEE conference on computer vision and pattern recognition*, 2009, pp. 1597–1604.

[25] Y.-F. Ma and H.-J. Zhang, "Contrast-based image attention analysis by using fuzzy growing," in *Proceedings of the eleventh ACM international conference on Multimedia*, 2003, pp. 374–381.

[26] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," *Adv. Neural Inf. Process. Syst.*, vol. 19, 2006.

[27] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, 1998.

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 385-399, November-2023
https://doi.org/10.47001/IRJIET/2023.711052

[28] S. Frintrop, M. Klodt, and E. Rome, "A real-time visual attention system using integral images," in *International Conference on Computer Vision Systems: Proceedings (2007)*, 2007.

[29] R. Achanta and S. Süsstrunk, "Saliency detection using maximum symmetric surround," in *2010 IEEE International Conference on Image Processing*, 2010, pp. 2653–2656.

[30] D. P. Bavirisetti and R. Dhuli, "Multi-focus image fusion using multi-scale image decomposition and saliency detection," *Ain Shams Eng. J.*, vol. 9, no. 4, pp. 1103–1117, 2018.

[31] S. Paul, I. S. Sevcenco, and P. Agathoklis, "Multi-exposure and multi-focus image fusion in gradient domain," *J. Circuits, Syst. Comput.*, vol. 25, no. 10, pp. 1–18, 2016, doi: 10.1142/S0218126616501231.

[32] Y. Wang and Y. Wang, "Fusion of 3-D medical image gradient domain based on detail-driven and directional structure tensor," *J. Xray. Sci. Technol.*, vol. 28, no. 5, pp. 1001–1016, 2020.

[33] R. Hong, C. Wang, Y. Ge, M. Wang, X. Wu, and R. Zhang, "Salience preserving multi-focus image fusion," in *2007 IEEE international conference on multimedia and expo*, 2007, pp. 1663–1666.

[34] G. Piella, "Image fusion for enhanced visualization: A variational approach," *Int. J. Comput. Vis.*, vol. 83, pp. 1–11, 2009.

[35] Y. Zhou, X. Yang, R. Zhang, K. Liu, M. Anisetti, and G. Jeon, "Gradient-based multi-focus image fusion method using convolution neural network," *Comput. Electr. Eng.*, vol. 92, p. 107174, 2021, doi: https://doi.org/10.1016/j.compeleceng.2021.107174.

[36] X. Ma, Z. Wang, and S. Hu, "Multi-focus image fusion based on multi-scale sparse representation," *J. Vis. Commun. Image Represent.*, vol. 81, p. 103328, 2021.

[37] H. Yin, Y. Li, Y. Chai, Z. Liu, and Z. Zhu, "A novel sparse-representation-based multi-focus image fusion approach," *Neurocomputing*, vol. 216, pp. 216–229, 2016.

[38] Y. Liu and Z. Wang, "Simultaneous image fusion and denoising with adaptive sparse representation," *IET Image Process.*, vol. 9, no. 5, pp. 347–357, 2015.

[39] B. Yang and S. Li, "Multifocus image fusion and restoration with sparse representation," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 4, pp. 884–892, 2009.

[40] B. Yang and S. Li, "Pixel-level image fusion with simultaneous orthogonal matching pursuit," *Inf. fusion*, vol. 13, no. 1, pp. 10–19, 2012.

[41] L. Chen, J. Li, and C. L. P. Chen, "Regional multifocus image fusion using sparse representation," *Opt. Express*, vol. 21, no. 4, pp. 5182–5197, 2013.

[42] X. Ma, S. Hu, S. Liu, J. Fang, and S. Xu, "Multi-focus image fusion based on joint sparse representation and optimum theory," *Signal Process. Image Commun.*, vol. 78, pp. 125–134, 2019.

[43] I. Zafar, E. A. Edirisinghe, and H. E. Bez, "Multi-exposure & multi-focus image fusion in transform domain," 2006.

[44] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "Multi-focus image fusion for visual sensor networks in DCT domain," *Comput. Electr. Eng.*, vol. 37, no. 5, pp. 789–797, 2011.

[45] Y. A. V. Phamila and R. Amutha, "Discrete Cosine Transform based fusion of multi-focus images for visual sensor networks," *Signal Processing*, vol. 95, pp. 161–170, 2014.

[46] J. Tang, "A contrast based image fusion technique in the DCT domain," *Digit. Signal Process.*, vol. 14, no. 3, pp. 218–226, 2004.

[47] M. A. Naji and A. Aghagolzadeh, "Multi-focus image fusion in DCT domain based on correlation coefficient," in *2015 2nd international conference on knowledge-based engineering and innovation (KBEI)*, 2015, pp. 632–639.

[48] M. A. Naji and A. Aghagolzadeh, "A new multi-focus image fusion technique based on variance in DCT domain," in *2015 2nd International conference on knowledge-based engineering and innovation (KBEI)*, 2015, pp. 478–484.

[49] Y. Liu, X. Chen, H. Peng, and Z. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Inf. Fusion*, vol. 36, pp. 191–207, 2017.

[50] J. Li *et al.*, "DRPL: Deep regression pair learning for multi-focus image fusion," *IEEE Trans. Image Process.*, vol. 29, pp. 4816–4831, 2020.

[51] Y. Zhang, Y. Liu, P. Sun, H. Yan, X. Zhao, and L. Zhang, "IFCNN: A general image fusion framework based on convolutional neural network," *Inf. Fusion*, vol. 54, pp. 99–118, 2020.

[52] R. Lai, Y. Li, J. Guan, and A. Xiong, "Multi-scale visual attention deep convolutional neural network for multi-focus image fusion," *IEEE Access*, vol. 7, pp. 114385–114399, 2019.

[53] M. Amin-Naji, A. Aghagolzadeh, and M. Ezoji, "Ensemble of CNN for multi-focus image fusion," *Inf. fusion*, vol. 51, pp. 201–214, 2019.

[54] M. Amin-Naji, A. Aghagolzadeh, and M. Ezoji, "CNNs hard voting for multi-focus image fusion," *J. Ambient Intell. Humaniz. Comput.*, vol. 11, pp. 1749–1769, 2020.

[55] H. Zhai and Y. Zhuang, "Multi-focus image fusion method using energy of Laplacian and a deep neural network," *Appl. Opt.*, vol. 59, no. 6, pp. 1684–1694, 2020.

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 385-399, November-2023
https://doi.org/10.47001/IRJIET/2023.711052

[56] V. Deshmukh, A. Khaparde, and S. Shaikh, "Multi-focus image fusion using deep belief network," in *Information and Communication Technology for Intelligent Systems (ICTIS 2017)-Volume 1 2*, 2018, pp. 233–241.

[57] F. Lahoud and S. Süsstrunk, "Fast and efficient zero-learning image fusion," *arXiv Prepr. arXiv1905.03590*, 2019.

[58] X. Zhang, "Deep Learning-Based Multi-Focus Image Fusion: A Survey and a Comparative Study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 9, pp. 4819–4838, 2022, doi: 10.1109/TPAMI.2021.3078906.

[59] H. Xu, J. Ma, Z. Le, J. Jiang, and X. Guo, "Fusiondn: A unified densely connected network for image fusion," in *Proceedings of the AAAI conference on artificial intelligence*, 2020, vol. 34, no. 07, pp. 12484–12491.

[60] B. Ma, Y. Zhu, X. Yin, X. Ban, H. Huang, and M. Mukeshimana, "Sesf-fuse: An unsupervised deep model for multi-focus image fusion," *Neural Comput. Appl.*, vol. 33, pp. 5793–5804, 2021.

[61] H. Zhang, Z. Le, Z. Shao, H. Xu, and J. Ma, "MFF-GAN: An unsupervised generative adversarial network with adaptive and gradient joint constraints for multi-focus image fusion," *Inf. Fusion*, vol. 66, pp. 40–53, 2021.

[62] H. Xu, J. Ma, J. Jiang, X. Guo, and H. Ling, "U2Fusion: A unified unsupervised image fusion network," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 1, pp. 502–518, 2020.

[63] H. Li, Y. Chai, and Z. Li, "A new fusion scheme for multifocus images based on focused pixels detection," *Mach. Vis. Appl.*, vol. 24, pp. 1167–1181, 2013.

[64] H. Li, X. Liu, Z. Yu, and Y. Zhang, "Performance improvement scheme of multifocus image fusion derived by difference images," *Signal Processing*, vol. 128, pp. 474–493, 2016.

[65] H. Li, H. Qiu, Z. Yu, and B. Li, "Multifocus image fusion via fixed window technique of multiscale images and non-local means filtering," *Signal Processing*, vol. 138, pp. 71–85, 2017.

[66] Q. Zhang, F. Wang, Y. Luo, and J. Han, "Exploring a unified low rank representation for multi-focus image fusion," *Pattern Recognit.*, vol. 113, p. 107752, 2021.

[67] K. Xu, Z. Qin, G. Wang, H. Zhang, K. Huang, and S. Ye, "Multi-focus image fusion using fully convolutional two-stream network for visual sensors," *KSII Trans. Internet Inf. Syst.*, vol. 12, no. 5, pp. 2253–2272, 2018.

[68] K. He, D. Zhou, X. Zhang, R. Nie, and X. Jin, "Multi-focus image fusion combining focus-region-level partition and pulse-coupled neural network," *Soft Comput.*, vol. 23, pp. 4685–4699, 2019.

[69] R. Hou, D. Zhou, R. Nie, and D. Liu, "Multi-focus color image fusion scheme using NSST and focus region detection," in *Proceedings of the 3rd international conference on multimedia and image processing*, 2018, pp. 7–11.

[70] D. M. Bulanon, T. F. Burks, and V. Alchanatis, "Image fusion of visible and thermal images for fruit detection," *Biosyst. Eng.*, vol. 103, no. 1, pp. 12–22, 2009.

[71] M. Hossny, S. Nahavandi, and D. Creighton, "Comments on'Information measure for performance of image fusion'," 2008.

[72] G. Cui, H. Feng, Z. Xu, Q. Li, and Y. Chen, "Detail preserved fusion of visible and infrared images using regional saliency extraction and multi-scale image decomposition," *Opt. Commun.*, vol. 341, pp. 199–209, 2015, doi: https://doi.org/10.1016/j.optcom.2014.12.032.

[73] B. Rajalingam, R. Priya, and R. Bhavani, "Hybrid Multimodal Medical Image Fusion Using Combination of Transform Techniques for Disease Analysis," *Procedia Comput. Sci.*, vol. 152, pp. 150–157, 2019, doi: https://doi.org/10.1016/j.procs.2019.05.037.

[74] O. R. Vincent and O. Folorunso, "A descriptive algorithm for sobel image edge detection," in *Proceedings of informing science & IT education conference (InSITE)*, 2009, vol. 40, pp. 97–107.

[75] C. S. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electron. Lett.*, vol. 36, no. 4, pp. 308–309, 2000.

[76] Y.-J. Rao, "In-fibre Bragg grating sensors," *Meas. Sci. Technol.*, vol. 8, no. 4, p. 355, 1997.

[77] A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance," *IEEE Trans. Commun.*, vol. 43, no. 12, pp. 2959–2965, 1995.

[78] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

[79] S. Li, R. Hong, and X. Wu, "A novel similarity based quality metric for image fusion," in *2008 International Conference on Audio, Language and Image Processing*, 2008, pp. 167–172.

[80] Y. Chen and R. S. Blum, "A new automated quality assessment algorithm for image fusion," *Image Vis. Comput.*, vol. 27, no. 10, pp. 1421–1432, 2009.

[81] Y. Han, Y. Cai, Y. Cao, and X. Xu, "A new image fusion performance metric based on visual information fidelity," *Inf. fusion*, vol. 14, no. 2, pp. 127–135, 2013.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 385-399, November-2023

https://doi.org/10.47001/IRJIET/2023.711052

Citation of this Article:

Karam Mohammed Ghazal, Dr. Ielaf O. Abdul Majjed Dahl, "Exploring Types of Multi-Focus Image Fusion" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 11, pp 385-399, November 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.711052

****

