

# Exploiting Superpixels for Multi-Focus Image Fusion

Areeba Ilyas <sup>1,†</sup> , Muhammad Shahid Farid <sup>1,\*,†</sup> , Muhammad Hassan Khan <sup>1,†</sup>  and Marcin Grzegorzec <sup>2,†</sup> 

<sup>1</sup> Punjab University College of Information Technology, University of the Punjab, Lahore 54000, Pakistan; areeba.ilyas@pucit.edu.pk (A.I.); hassankhan@pucit.edu.pk (M.H.K.)

<sup>2</sup> Institute of Medical Informatics, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany; grzegorzec@imi.uni-luebeck.de

\* Correspondence: shahid@pucit.edu.pk

† These authors contributed equally to this work.

**Abstract:** Multi-focus image fusion is the process of combining focused regions of two or more images to obtain a single all-in-focus image. It is an important research area because a fused image is of high quality and contains more details than the source images. This makes it useful for numerous applications in image enhancement, remote sensing, object recognition, medical imaging, etc. This paper presents a novel multi-focus image fusion algorithm that proposes to group the local connected pixels with similar colors and patterns, usually referred to as superpixels, and use them to separate the focused and de-focused regions of an image. We note that these superpixels are more expressive than individual pixels, and they carry more distinctive statistical properties when compared with other superpixels. The statistical properties of superpixels are analyzed to categorize the pixels as focused or de-focused and to estimate a focus map. A spatial consistency constraint is ensured on the initial focus map to obtain a refined map, which is used in the fusion rule to obtain a single all-in-focus image. Qualitative and quantitative evaluations are performed to assess the performance of the proposed method on a benchmark multi-focus image fusion dataset. The results show that our method produces better quality fused images than existing image fusion techniques.



**Citation:** Ilyas, A.; Farid, M.S.; Khan, M.H.; Grzegorzec, M. Exploiting Superpixels for Multi-Focus Image Fusion. *Entropy* **2021**, *23*, 247. <https://doi.org/10.3390/e23020247>

**Keywords:** multi-focus image fusion; image enhancement; information fusion color distance models for fusion

Academic Editor: Jiayi Ma

Received: 29 January 2021

Accepted: 17 February 2021

Published: 21 February 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

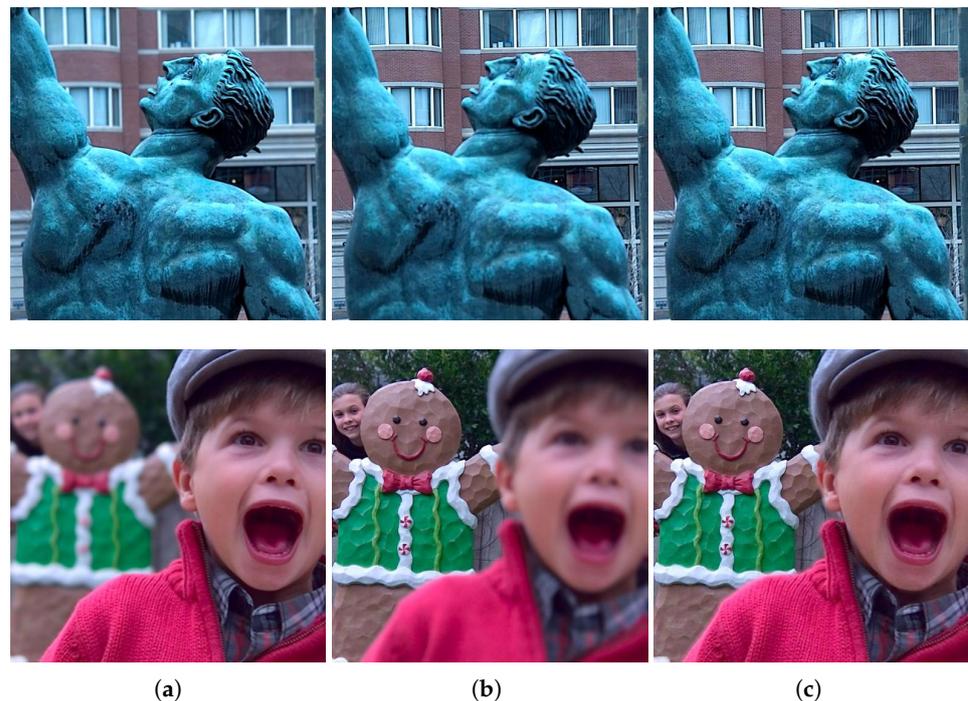
## 1. Introduction

Due to limited depth-of-field (DOF), it is not easy for cameras to capture an image that has all focused objects. Usually, in a camera captured image, the objects that lie in the range of the depth-of-field are sharp and the remaining objects tend to be blurred [1]. An additional cause of image blurring includes atmospheric scatter and spatial and temporal sensor integration [2]. Images that are partially focused are not enough for sound understanding and obtaining accurate results in different applications of computer vision such as object recognition and extraction [3], remote sensing and surveillance [4], image enhancement [5], medical imaging [6], etc. To resolve this issue, multi-focus image fusion algorithms are proposed in which a fused image of an extended depth-of-field is constructed by integrating the additional information of multiple images of the same scene.

In recent years, numerous multi-focus image fusion methods have been proposed. Some of them directly operate on image pixels and regions to obtain an all-in-focus image while others exploit the transform domains to achieve this task. The former methods are known as spatial domain based fusion methods, and the latter are categorized as transform domain based methods [7,8]. In transform domain based fusion, a fused image is regenerated from fused transform coefficients of the source images. The transform domain based fusion methods perform well for separating the discontinuities at edges, but do not give the representation for line and curve discontinuities [9]. They are computationally expensive techniques and may produce undesirable artifacts in fused image [10].

In spatial domain based fusion, the rules of fusion are applied to pixels or a region of input images directly without conversion to other representations to generate the fused image. Most spatial domain fusion algorithms operate at the pixel level, e.g., [11–14], and produce a fused image by weighting the corresponding pixels of source images. Some spatial domain based fusion techniques work at the region level, e.g., [15–17]. These methods compute focus measures for each region and construct a fused image by selecting a region with more clarity. The spatial domain based fusion algorithms provide high spatial resolution. However, the results might suffer from unwanted artifacts, e.g., the ghosting effect, which mainly occurs due to incorrect categorization of the pixels or regions as focused or de-focused. To overcome these problems, some techniques in the spatial domain use real values instead of binary values as focus measures and use the weighted averaging of pixel values to obtain the fused image [18,19].

In this paper, we propose a novel spatial domain based multi-focus image fusion algorithm. The proposed method is built on the idea of grouping the neighboring pixels with similar colors and patterns into larger pixels (also known as superpixels) to obtain an accurate decision map. We observe that such a grouping helps reduce the false categorization of pixels as focused and de-focused. The statistical properties of the computed superpixels are analyzed to distinguish the focused and de-focused pixels, generating the initial focus map, which is further processed to obtain a more accurate final decision map. Different tests performed to evaluate the performance of the proposed method reveal its effectiveness. A few of the sample fusion results achieved by the proposed method are shown in Figure 1.



**Figure 1.** Results of the proposed algorithm on sample multi-focus images from the test dataset. (a) Source images with foreground focused, (b) source images with background focused, and (c) fused images using the proposed method.

The rest of this paper is organized as follows: A brief review of the literature on multi-focus image fusion is presented in Section 2. The proposed MIFalgorithm is presented in Section 3. The experimental evaluation through different qualitative and quantitative tools is carried out in Section 4. The performance of the proposed algorithm with different color distance models and the impact of other parameters is analyzed in Section 5. The conclusions are drawn in Section 6.

## 2. Background and Literature Review

The main step in any multi-focus image fusion algorithm is the detection of the focused region in the source multi-focus images to obtain a so-called decision map, also called the focus map. The source images are then fused to obtain an all-in-focus image. Based on the representation and working of the focus map estimation and fusion process, the existing MIF algorithms can be categorized into four groups: multi-scale transform based methods, feature space transform based methods, spatial domain based methods, and neural networks based methods [8,20].

The multi-scale transform based methods generally split the input images into a multi-scale domain and then fuse the transformed coefficients using some fusion rule. The all-in-focus image is generated from these fused coefficients [21]. Multi-scale transform methods have a wide range of related techniques sequentially implemented in image fusion like pyramid decomposition [22], wavelet transform [23,24], curvelet transform (CVT) [25], and contourlet transform [26,27]. In the MIF algorithm presented in [28], the source images are split into approximation and detail coefficients at different levels, and the coefficients are fused by applying several fusion rules. Various level fused coefficients are merged to obtain the desired all-in-focus image. Sheng et al. [29] proposed a support value transform based method that uses a support vector machine to achieve the fused image. The discrete cosine transform has also been exploited for image fusion, e.g., [30,31]

The feature space transform based fusion has emerged as another popular mean for multi-focus image fusion [32]. These fusion methods estimate the focus map through a single scale feature space of source images. References [1,13,33,34] are a few examples of such MIF methods. The sparse representation based fusion method in [33] divides the input images into patches known as sparse coefficients using a dictionary. The coefficients are then combined, and the fused image is generated. The fusion method in [1] uses the feature space from a robust principal component analysis technique for decomposition, and local sparse features are used to compute the all-in-focus image. The higher order singular-value decomposition based method [34] picks informative patches from input images by estimating the activity level of each patch. This information helps to obtain the fused image. Another higher order singular-value decomposition based method was presented in [35]. Multi-focus image fusion with the dense scale-invariant feature transform method (DSIFT) [13] uses dense descriptors to calculate the activity level in the patches of source images to form an initial decision map. After refining and feature matching of the initial map, it is used to form the fused image.

In spatial domain based fusion methods, fused images are computed by processing the pixels of the input multi-focus images. Some methods operate at the pixel level, e.g., [11,14,36]; some are block based, e.g., [37–39]; and others are region based methods, such as [40–42]. Usually, pixel-by-pixel averaging of source images does not achieve plausible results as it may introduce different artifacts in the fused image [13,43]. Therefore, the region based and block based spatial domain fusion methods have been introduced. In these methods, the input images are divided into blocks or segmented regions, and the sharpness of each block is calculated to form the focus map [44].

Image segmentation is a difficult task, so it is hard to maintain the quality of fused images obtained by region based methods. However, as an image block may contain both focused and defocused areas, the fused image computed by the block based method may exhibit anisotropic blur and misregistration. To resolve this problem, multi-scale image fusion methods have been introduced. The method in [45] is based on a multi-scale structure based focus measure. In this method, a binary-scale scheme is used, which calculates the weighted gradient of a focused region by applying small-scale focus measures. It proves helpful in reducing the anisotropic blur and misregistration with large-scale focus measures.

In recent years, many pixel based spatial domain methods have been proposed. These methods include image matting based [18] and guided filtering based [46] methods, which compute the fused image by extracting enough information from the source images and

achieve high efficiency by preserving spatial consistency. The MIF method proposed in [18] focuses on obtaining all-in-one focus images from detected focused regions in multi-focus images by using morphological filtering and the image matting technique. In the guided filtering based method [46], two-scale image decomposition is done to produce focused images from the base layer and detail layer.

Some pixel based methods, such as [47–49], solve the maximum optimization problem by estimating a weight map. The algorithm in [47] has the benefit of a multi-spectral algorithm along with spatial domain methods. The combined activity estimation of high and low frequencies of source images is calculated, and a new smoothness term is introduced for the alignment of the solution to the boundaries of the defocused and focused pixel. In [48], multi-focus image fusion was proposed using edge and multi-matting models. A multi-focus image fusion algorithm based on random walks [49] analyzes the feature area locally and identifies nodes globally to compute connected graphs of random walks.

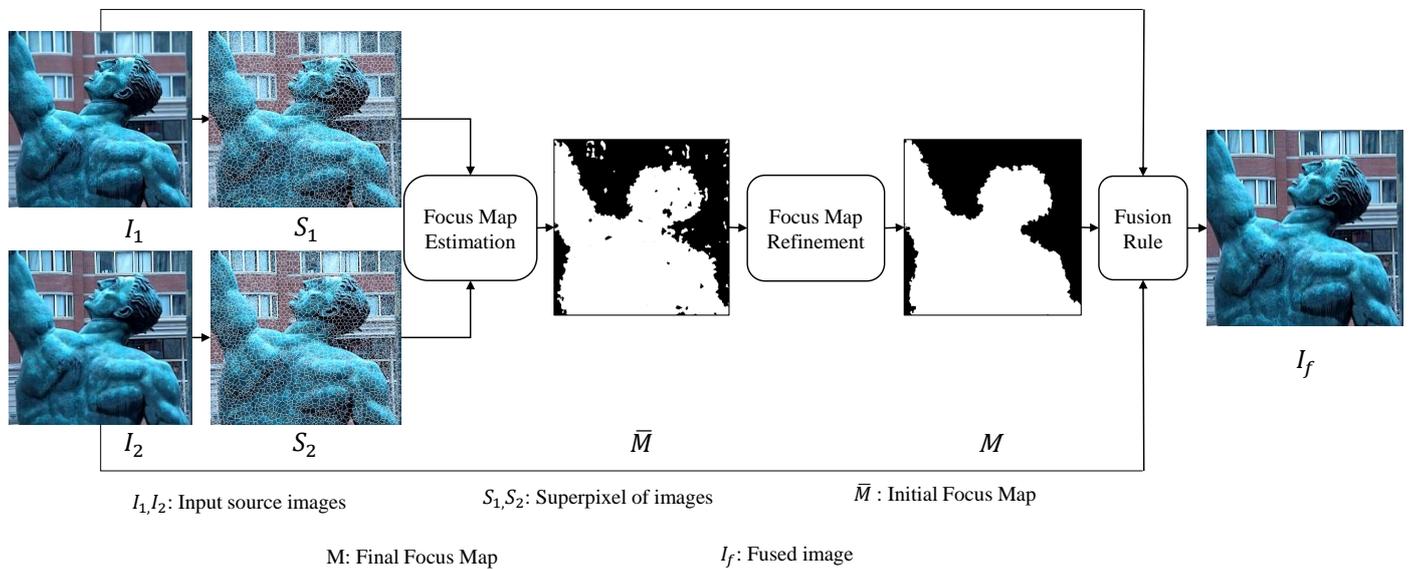
Multi-focus image fusion has also been achieved through deep neural networks in several ways; the majority of these techniques rely on the detection of the focused region [50]. In the fusion method presented in [50], features are extracted through a u-shaped network to obtain high- and low-level frequency texture information. It directly maps multi-focus images to fused images instead of detecting focused regions. The pulse coupled neural network (PCNN) is also explored for image fusion, e.g., [16]. In these methods, the source multi-focus images are decomposed into fixed-size blocks, and the image Laplacian of each block is computed to obtain the feature maps. These feature maps are used as the input of the PCNN as in [16].

Recently, deep learning has also been explored for efficient image fusion [51]. A number multi-focus image fusion techniques based on deep learning (DL) have been presented using convolutional neural networks (CNNs), such as in [52–56]. The method in [52] proposes a pixel-wise convolution neural network for image fusion, and the deep convolutional neural network (CNN) based method [54] uses the Siamese network for feature extraction and activity level measurement in source images to obtain a focus map. The CNN based MIF method presented in [53] also uses Siamese multi-scale feature extraction. In [55], a multi-level dense network was presented for fusing multi-focus images. It extracts shallow and dense features from images from a mixture of many distributions. The convolutional sparse representation (CSR) is used for multi-focus image fusion, e.g., [57–59]. The algorithm in [59] uses Taylor expansion and convolutional sparse representation for fusing multi-focus images. Morphological component analysis and convolutional sparse representation were used in [58] to obtain texture features using pre-learned dictionaries. The obtained sparse coefficients of the source images were merged, and a fused image was obtained. Generative adversarial networks (GANs) have also been exploited for increasing the depth of field, e.g., [60–62].

### 3. Proposed Method

We propose a spatial domain based multi-focus image fusion algorithm that builds on the idea that pixels in a small locality in an image are highly autocorrelated—having similar colors and textures—and thus can be merged into larger patches known as superpixels to improve the speed and quality of different tasks. These superpixels are more expressive than single pixels as they carry more distinctive statistical properties when compared with other superpixels. Moreover, we observe that they have a greater tendency to be distinguished as focused and out-of-focused compared to single pixels. We exploit this important characteristic to design a novel and efficient image fusion algorithm. The proposed method works in three steps. First, superpixels in the source multi-focus images are computed. Second, using the statistical properties of superpixels, they are categorized as focused and de-focused, from which initial focus maps are constructed. The initial focus map is refined using a spatial consistency constraint. Third, the refined focus map is used to obtain the fused image from source multi-focus images with the help of a fusion rule. A

schematic diagram of the proposed method is shown in Figure 2. These steps are described in detail in the following sections.



**Figure 2.** Schematic diagram of the proposed multi-focus image fusion algorithm.

### 3.1. Superpixels Computation

Several algorithms have been proposed to obtain superpixels from an image, e.g., [63–65]. Like most existing superpixel algorithms, the proposed technique also uses pixel color and its spatial position in deciding their merging. Our method extends the simple linear iterative clustering (SLIC) algorithm [65], which uses  $k$ -means clustering to generate superpixels, to obtain superpixels better suited for image fusion. In particular, similar to [65], we divide the image pixels into  $k$  number of clusters based on their color and location. Superpixels are computed based on the mean color values of pixels. For an image with  $N$  pixels, the size of each superpixel is approximately  $N/k$  pixels. Regular spaced grid pixels  $S$  are computed where  $S = \sqrt{N/K}$ . Every initial cluster center is then sampled based on  $S$  to produce equally sized superpixels. After the comparison, these centers are then moved forward to the lowest gradient position. Then, each pixel of the image is associated with the nearest cluster center through a distance measure  $D_s$ .  $D_s$  is computed considering the pixel spatial distance  $d_s$  and the color difference  $d_c$ .

The spatial distance of the pixel located at  $(x_i, y_i)$  and a cluster center  $(x_k, y_k)$  is computed using the Euclidean distance metric.

$$d_s = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \tag{1}$$

For color distance measure, unlike the method in [65], we use the CIEDE2000 color difference model, which produces better superpixels for image fusion. We performed numerous experiments and tested different color distance models to evaluate their effectiveness for pixel focus detection, which favored the CIEDE2000 metric. It is a complicated, but accurate color distance measure [66]. In this model, the color difference between two pixels with color in the  $L^*a^*b^*$  model  $(L_1^*a_1^*b_1^*)$  and  $(L_2^*a_2^*b_2^*)$  is computed as:

$$d_c = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2} + R_T \frac{\Delta C'}{k_C S_C} \frac{\Delta H'}{k_H S_H} \tag{2}$$

where  $\Delta L'$  is the difference in luminance components and  $\Delta C'$  and  $\Delta H'$  are computed from the chroma components  $a^*, b^*$ .

$$\begin{aligned}\Delta L' &= L_2^* - L_1^* \\ \Delta C' &= C_2' - C_1'\end{aligned}$$

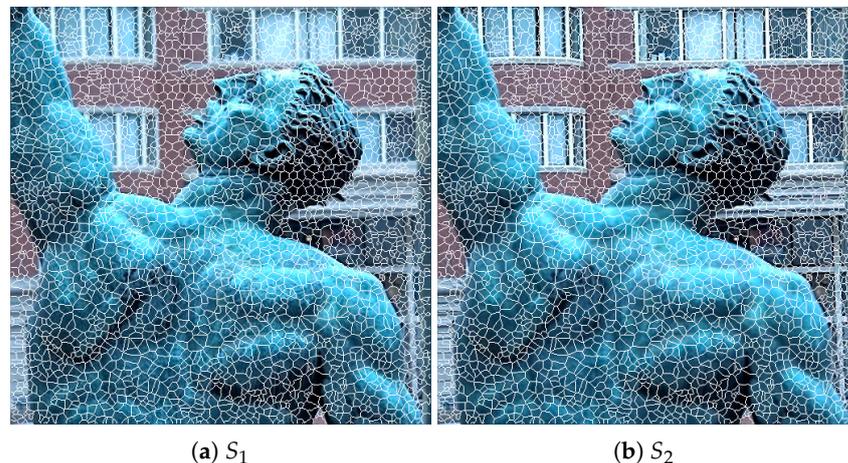
with  $C_1' = \sqrt{(a_1^*)^2 + (b_1^*)^2}$ ,  $C_2' = \sqrt{(a_2^*)^2 + (b_2^*)^2}$  and  $\Delta H' = \sqrt{C_1' C_2'} \sin(\Delta h' / 2)$ .

The weighting factor  $K_L$ ,  $K_C$ , and  $K_H$  in (2) are usually set to unity.  $S_L$ ,  $S_C$ , and  $S_H$  are the compensation of lightness, chroma, and hue. To keep the discussion simple, the complete details of the metric are not presented here, and they can be found in [66].

Finally, the two distance measures are combined as in [65] to obtain  $D_s$ . It is the sum of the spatial distances (1) and color distances (2), which are normalized by grid interval  $S$ . The factor  $m$  in (3) is used to control the compactness of the superpixels.

$$D_s = d_c + \frac{m}{S} d_s \quad (3)$$

Let  $I_1$  and  $I_2$  be two source multi-focus images. In the rest of the text, they are simply referred to as source images. Using the proposed algorithm, the superpixels in  $I_1$  are computed, which are mapped to  $I_2$  to divide  $I_2$  into the corresponding set of superpixels. That is, the superpixels in  $I_2$  are not computed; instead, the superpixel labels computed in  $I_1$  are assumed for  $I_2$  as well, and the values of their centers are updated with respect to  $I_2$ . This helps to establish the correspondence between the superpixels in  $I_1$  and  $I_2$ . We recall that both images are complementary and represent the same scene, but with different focused regions. Therefore, the superpixel structure of the two images is kept the same so that the correspondence between them can be established. Figure 3 shows the superpixels computed using the proposed strategy on a sample multi-focus image pair (shown in Figure 1).



**Figure 3.** Superpixel computation using the proposed strategy. (a) Superpixels of  $I_1$  (foreground focused). (b) Superpixels of  $I_2$  (background focused).

### 3.2. Focus Map Estimation

A superpixel carries various statistical properties, e.g., mean, standard deviation, variance, and kurtosis, which can be used to draw different characteristics of the superpixel. The standard deviation is considered to be an important and widely used index to measure information diversity and variability in an image [67,68]. It has been shown in different studies, e.g., [69], that the variance (or standard deviation) of the de-focused regions is generally less than the sharp regions. In the proposed method, we exploit this property to designate the superpixels as focused and de-focused. We calculate the standard deviation

of each superpixel in the source images and use it to estimate the focus map. We found that the standard deviation of  $L$  channel serves as a better discriminant than the standard deviation computed over all three channels, i.e.,  $l$ ,  $a$ , and  $b$ . Therefore, for each superpixel,  $s$ , the standard deviation  $\sigma$  of the  $L$  values of its constituent pixels is computed.

$$\sigma_{1,i} = \sqrt{\frac{1}{|s_{1,i}|} \sum_{(x,y) \in s_{1,i}} (I_1(x,y) - \mu_{1,i})^2} \quad (4)$$

where  $s_{1,i}$  is the  $i$ -th superpixel in image  $I_1$ ,  $|s_{1,i}|$  is the number of pixels in  $s_{1,i}$ , and  $\mu_{1,i}$  is the mean of the pixels in superpixel  $s_{1,i}$ .

$$\mu_{1,i} = \frac{1}{|s_i|} \sum_{(x,y) \in s_i} I_1(x,y)$$

The standard deviation of each superpixel in  $I_2$ ,  $\sigma_{2,i}$ , is computed analogously. The standard deviations of the corresponding superpixels in  $I_1$  and  $I_2$  are compared, and the focus map  $\bar{M}$  is constructed. If  $\sigma_{1,i}$  is greater than  $\sigma_{2,i}$ , this means  $s_{1,i}$  is focused in  $I_1$  and the same region  $s_{2,i}$  is de-focused in  $I_2$ . In the focus map  $\bar{M}$ , the focus value of pixels corresponding to  $s_{1,i}$  is set to one, and it is set to zero if  $\sigma_{1,i} \not> \sigma_{2,i}$ . That is,

$$\bar{M}(x,y) = \begin{cases} 1 & \text{if } \sigma_{1,i} > \sigma_{2,i} \\ 0 & \text{if } \sigma_{1,i} \leq \sigma_{2,i} \end{cases} \quad (5)$$

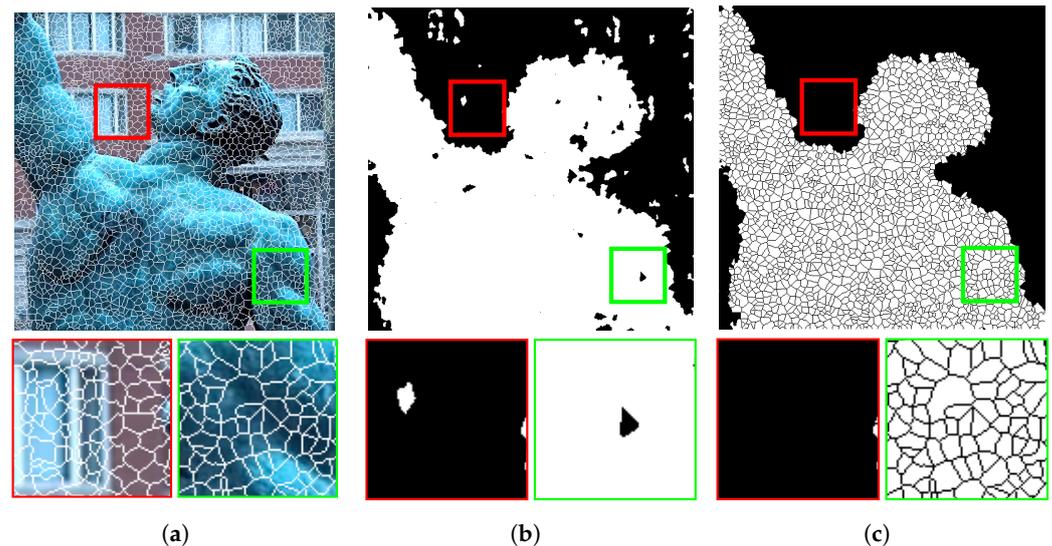
Figure 4 shows the focus map constructed from the superpixels shown in Figure 3 using Equation (5).

It is possible that in the initial focus map obtained through the proposed algorithm, a few superpixels are incorrectly categorized, as can be noted in Figure 4. Some focused superpixels are falsely marked as de-focused, and a few de-focused patches are incorrectly marked as focused. These false positives and false negatives should be resolved to obtain a high-quality fused image. To this end, a neighbor constraint technique is proposed. We observed from experiments that the incorrectly categorized superpixels usually appear as a single entity, and most of its neighbors are correctly identified. Based on this observation, we form a neighbor constraint rule to eliminate the false positives and negatives from the map. The rule examines each superpixel in the map, investigates all its neighboring superpixels, and assigns the category that most of its neighbors have. Specifically, it is marked as focused if more than half of its neighbors are marked as focused, and vice versa.



Figure 4. Initial focus map generated using the proposed strategy for the images shown in Figure 3.

Figure 5 shows the results of the proposed refinement process. Figure 5a shows a source image with superpixels overlaid on it; two sample focused and de-focused regions magnified to ease the inspection are shown below it. Figure 5b shows the initial focus map; superpixels structures are also overlaid to highlight the details, and selected regions are also shown below this figure. The results of the refinement procedure are shown in Figure 5c. The results show that most incorrectly marked regions were successfully recovered by the proposed refinement strategy.

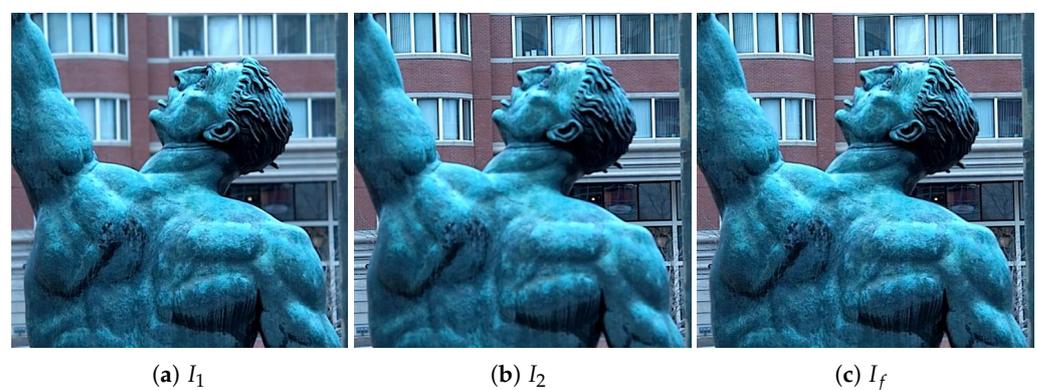


**Figure 5.** Initial focus map refinement: (a) a sample multi-focus image with superpixels structures overlaid on it, (b) the obtained focus map, and (c) the refined focus map.

### 3.3. Fusion Rule

The final focus map  $M$  obtained after refinement is used to fuse the contents of the source images  $I_1$  and  $I_2$  to generate the all-in-focus image  $I_f$ . The fusion rule simply takes the pixel value from  $I_1$  if the focus map value is one and from  $I_2$  otherwise. Figure 6 shows the fused image generated from the source images shown in Figure 3.

$$I_f(x, y) = \begin{cases} I_1(x, y) & \text{if } M(x, y) = 1 \\ I_2(x, y) & \text{if } M(x, y) = 0 \end{cases} \quad (6)$$



**Figure 6.** Fusion results using the refined focus map: (a) first source image  $I_1$ , (b) second source image  $I_2$ , and (c) fused image  $I_f$ .

#### 4. Experiments and Results

In this section, we evaluate the performance of the proposed multi-focus image fusion method and compare it with the existing MIF techniques. The performance was assessed both qualitatively and quantitatively, and various analyses were also performed to evaluate the robustness of the proposed method. The experiments were performed on the recent widely used Lytro multi-focus dataset [43,70]. The dataset contains 20 color multi-focused image pairs, each of size  $520 \times 520$ . Sample image pairs from the dataset are shown in Figure 7.

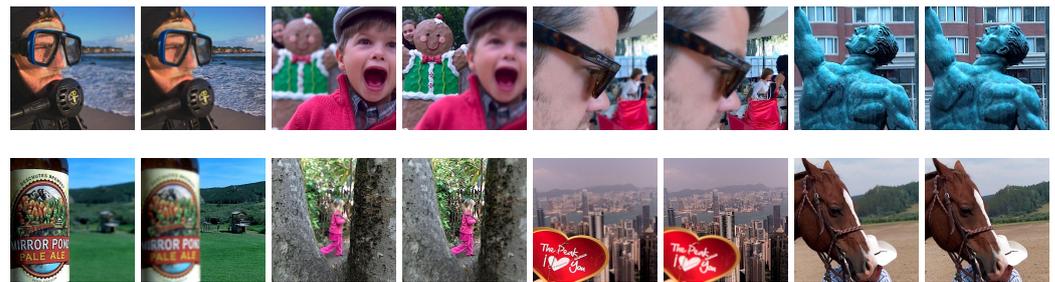


Figure 7. Sample multi-focus image pairs from the Lytro dataset.

##### 4.1. Compared Methods

To evaluate the effectiveness of the proposed method, its performance was compared with existing well-known multi-focus image fusion (MIF) algorithms. In particular, nine representative methods were selected for the comparison. The discrete cosine harmonic wavelet transform (DCHWT) based method [24] reduces the fusion complexity by using the discrete cosine transform (DCT) and discrete harmonic wavelet coefficients. The wavelet based statistical sharpness measure (WSSM) [71] utilizes the spreading of the wavelet coefficient distribution as a sharpness measure and evaluates it by the adaptive Laplacian mixture model. The pulse coupled neural network (PCNN) method [26,72] uses the orientation information to obtain the fused image. The source images are decomposed into fixed-size blocks, and the image Laplacian of each block is computed to obtain feature maps that are used in PCNN for fusion. The DCTL method [31] uses DCT coefficients and Laplacian pyramids to achieve fusion. The method presented in [73] uses Independent Component Analysis (ICA) and topographic independent component analysis bases for fusion.

In image fusion using luminance gradients (IFGD) [74], the highest gradients of each pixels' locations of the source images are blended, and then, fused luminance is obtained by the reconstruction technique. In the nonsubsampling contourlet transform (NSCT) based MIF algorithm [27], NSCT decomposition is used to obtain the selection rules for sub-band coefficients. In the PCA method [75], the fused image is obtained using wavelets and principal component analysis (PCA). The CSR fusion algorithm [76] decomposes each source image into base and detail layers with the help of convolution sparse representation. It performs different fusion techniques on both layers and then merges the results. A deep convolutional neural network (CNN) based MIF method was presented in [54]. In particular, it uses the Siamese network for feature extraction and activity level measurement in source images.

##### 4.2. Qualitative Performance Evaluation

The qualitative performance analysis was performed by analyzing the quality of fused images visually. For the visual inspection, we present the fusion results obtained by the compared methods and our algorithm. For comparison, "Swimmer" and "Cookie" image pairs are selected from the test dataset. The visual inspection is difficult to perform and generalize as each multi-focus image pair has different information about focused and

de-focused regions. Nevertheless, it can certainly provide some insight into the quality of fused images obtained from different algorithms.

Figure 8 presents the fused images of the “Swimmer” image pair obtained by our algorithm and the compared methods. One can note that the results of the DCTLP method suffer from undesirable ringing artifacts: the fused image has lost most of the sharpness around the strong edges. Some blurred regions can be spotted in the results of the PCNN and NSCT methods. The fused image obtained by IFGD has lost some fine details due to extra brightness. The ghosting and ringing artifacts can be seen in the results of the PCA and NSCT algorithms as well. The ghosting artifact is also visible around the edges in the fused image produced by the ICA method. The fused images produced by WSSM, CNN, CSR, and our method are of good quality and are hard to visually differentiate. Similar observations can be made from the visual inspection of the results of ours and the compared methods on the “Cookie” image pair shown in Figure 9.



**Figure 8.** Visual quality assessment of the fusion results achieved by the proposed and the compared methods on the “Swimmer” multi-focus image pair from the test dataset. WSSM, wavelet based statistical sharpness measure; PCNN, pulse coupled neural network; DCTLP, discrete cosine transform Laplacian pyramid; IFGD, image fusion using luminance gradients; CSR, convolutional sparse representation.



**Figure 9.** Visual quality assessment of the fusion results achieved by the proposed and the compared methods on the “Cookie” multi-focus image pair from the test dataset.

The visual quality inspection shows that the results produced by our method are crisp and are free of ghosting and ringing artifacts. A qualitative analysis, however, is difficult to perform due to various reasons including viewing conditions, trained subjects, etc., and it is particularly laborious on large datasets. Therefore, to truly assess the performance of an algorithm, along with the visual inspection, an objective fusion quality assessment is necessary.

#### 4.3. Quantitative Performance Evaluation

The objective quality assessment of fused images is a tough task as the ground truth images of multi-focus images cannot be acquired. To objectively evaluate the performance of MIF algorithms, many metrics have been introduced, but none of them are certainly better than the others. Therefore, to evaluate the performance of fusion methods, it is desirable to use multiple metrics. These metrics evaluate the quality of the fused image with reference to the source multi-focus images. To perform an extensive objective performance evaluation of the proposed method and other competing techniques, we chose 12 fusion quality assessment metrics and computed the results over the whole dataset. These metrics assess the quality of a fused image using its different characteristics and modalities. Based on these characteristics, the metrics are generally grouped into four categories [8,77]: information theory based metrics, image structural similarity based metrics, image feature based metrics, and human perception based metrics. These metrics are summarized in Table 1.

**Table 1.** Objective image fusion quality assessment metrics used in the performance evaluation.

Metric	Metric Description and Reference
Information theory based metrics	
$NMI$	Normalized mutual information [78,79]
$Q_{NCIE}$	Non-linear correlation metric [80]
$VIFF$	Visual information fidelity metric [81]
Feature based metrics	
$Q^{AB/F}$	Gradient based metric [82]
$Q_G$	Gradient based metric [83]
$Q_{SF}$	Spatial frequency based metric [84]
$Q_M$	Multi-scale scheme metric [85]
$Q_P$	Phase congruency based metric [86]
Structural similarity based metrics	
$Q_S$	Piella's metric [87]
$Q_C$	Cvejie's metric [88]
$Q_Y$	Yang's metric [89]
Human perception based metrics	
$Q_{CB}$	Chen Blum's metric [90]

- Information theory based metrics analyze the quality of the fused images by checking the amount of information transferred from the source images to the fused image. These metrics include: normalized mutual information ( $NMI$ ) [79], visual information fidelity ( $VIFF$ ) [81], and the nonlinear correlation information entropy metric ( $Q_{NCIE}$ ) [80].
- Image feature based metrics calculate the quality of fused image by analyzing the transferred features from the source images to the fused image. Gradient based fusion performance metrics  $Q^{AB/F}$  [82], ( $Q_G$ ) [83], image fusion metric based on multiscale scheme ( $Q_M$ ) [85], image fusion metric based on spatial frequency ( $Q_{SF}$ ) [84], and image fusion metric based on phase congruency ( $Q_P$ ) [86,91] are examples of these metrics.
- Image structural similarity based metrics: The measurement of image similarity depends on the proof that the human visual system is profoundly adjusted to structural information, which is exploited in these metrics to assess the quality of a fused image. The loss of structural information is very important to estimate the image distortion. These metrics include Piella's metric [87], Cvejie's image fusion metric ( $Q_C$ ) [88], and Yang's image fusion metric ( $Q_Y$ ) [89].
- Human perception inspired fusion metrics, e.g., Chen Blum metric  $Q_{CB}$  [90], measure the quality of fused images on the basis of the saliency map, which is generated by filtering and calculating the preservation of contrast.

The objective performance analysis of the proposed and the compared methods was carried out using all twelve metrics described in Table 1. Each metric's scores was computed over the whole dataset using the implementation presented in [77], and the results were averaged to obtain an overall quality score. The comparison is presented in Table 2, and the best performing algorithm on each metric is marked in bold. The results show that for the widely used  $Q^{AB/F}$  metric, our method performed the best, achieving the highest score 0.7539. The proposed method outperformed all competing techniques in most metrics including  $Q^{AB/F}$ ,  $Q_G$ ,  $Q_M$ ,  $Q_{SF}$ ,  $Q_Y$ , and  $Q_C$ . The performance of our method in the rest of the metrics was also appreciative, where it was the second or the third best algorithm. These statistics reveal that the proposed algorithm overall produces better quality fused images than most compared methods. Both qualitative and quantitative analyses reflect the effectiveness of the proposed method.

**Table 2.** Objective performance evaluation of the proposed and the compared methods on the Lytro dataset. The best scores for each metric are highlighted in bold.

Metrics	DCHWT	WSSM	PCNN	DCTLP	IFGD	ICA	NSCT	PCA	CNN	CSR	Proposed
$Q^{AB/F}$	0.7212	0.7305	0.7036	0.6526	0.7174	0.7445	0.2314	0.5992	0.7514	0.7366	<b>0.7539</b>
VIFF	0.9021	0.9333	0.8565	0.8388	<b>1.0456</b>	0.9386	0.5851	0.7809	0.9548	0.9329	0.9344
NMI	0.9176	0.9748	<b>1.2068</b>	0.8296	0.5223	0.9374	0.5645	0.8939	1.0725	0.9921	1.1972
$Q_G$	0.6153	0.6758	0.6745	0.5390	0.6181	0.6787	0.2070	0.5340	0.7029	0.6447	<b>0.7291</b>
$Q_{NCIE}$	0.8291	0.8321	<b>0.8497</b>	0.8252	0.8145	0.8301	0.8159	0.8277	0.8384	0.8341	0.8472
$Q_M$	0.8821	0.9331	2.2555	0.6983	0.5633	1.0264	0.1796	0.4820	2.3431	1.3400	<b>2.5074</b>
$Q_{SF}$	−0.0900	−0.0899	−0.1050	−0.0941	−0.0744	−0.0684	−0.1522	−0.3890	−0.0321	−0.0490	− <b>0.0314</b>
$Q_P$	0.7839	0.8201	0.7483	0.6892	0.7632	0.8197	0.0486	0.7500	<b>0.8428</b>	0.8283	0.8289
$Q_S$	0.9477	0.9438	0.9133	0.9298	0.8901	<b>0.9547</b>	0.5802	0.9217	0.9485	0.9467	0.9463
$Q_C$	0.8019	0.8097	0.8183	0.7698	0.7485	<b>0.8334</b>	0.3853	0.7978	0.8253	0.7980	0.8258
$Q_Y$	0.9219	0.9567	0.9672	0.8747	0.8526	0.9515	0.3792	0.8490	0.9653	0.9340	<b>0.9838</b>
$Q_{CB}$	0.6977	0.7887	0.7476	0.6185	0.6118	0.7130	0.4778	0.6325	0.7816	0.7638	<b>0.8054</b>

#### 4.4. Computational Complexity

We also analyzed the computational complexity of the proposed and the compared methods. All MIF algorithms were executed over the whole dataset, and their average execution time was computed and is reported in Table 3. These times also include the file I/O time, if any, and were computed on an Intel® Core™ i7 2.5 GHz processor with 8 GB RAM and 64 bit operating system. The results show that the PCA algorithm is the fastest among all approaches; the proposed method takes an average 122.61 seconds to generate a fused image. The main contributor to this time is the complex computation involved in estimating the superpixels and using the CIEDE2000 color distance model. An efficient implementation can help reduce its computational complexity.

**Table 3.** Execution time (seconds) comparison of the proposed and the compared MIF algorithms.

Method	DCHWT	WSSM	PCNN	DCTLP	IFGD	ICA	NSCT	PCA	CNN	CSR	Proposed
Time	9.39	215.49	1.31	0.33	1.01	11.30	174.00	0.04	106.32	345.76	122.61

#### 4.5. Summary

The performance of the proposed MIF method was evaluated using different tools and techniques. To appreciate the performance of the proposed method, visual results achieved by the proposed and the compared MIF algorithms were presented in the preceding sections. From the visual inspection of the results presented in Figures 8 and 9, we see that the results of most compared methods suffer from different impairments such as the ghosting effect, ringing artifacts, and blurry regions. In contrast, the results of the presented methods are crisp and free from these artifacts. The objective performance evaluation metrics confirm this. The advantages of the proposed method lie in its simplicity and ability to compute an accurate focus map. Unlike other competing methods that employ complex mechanisms to estimate the focus map and fusion strategies, the proposed method uses a simple fusion rule. We recall that the proposed MIF algorithm is built on the fact that the pixels in a small locality in an image are highly autocorrelated—having similar colors and textures. Based on this phenomenon, the pixels in the image are merged into larger patches, which provide the grounds for an accurate focus map, resulting in good fusion results. One limitation of the present proposed method is its inability to work with grayscale images. Since pixel color is an important clue in our focus map estimation method, its absence can degenerate its performance. However, extending it to grayscale images would certainly be an interesting investigation that we plan to carry out in the future.

## 5. Performance Analysis Using Different Color Distance Models and the Impact of other Parameters

In this section, we report different experiments and analyses that we performed to investigate the performance of the proposed image fusion algorithm. We performed three sets of experiments, first to evaluate the different color distance models for efficient superpixel estimation for fusion. The second set of experiments analyzed the impact of the number of superpixels on the performance of the proposed algorithm. In the third set of experiments, the contribution of color distance and spatial distance towards the overall distance measure was analyzed, and their impact on the performance of the proposed method was investigated.

### 5.1. Fusion Performance with Different Color Distance Models

We recall that determining the clusters to which the pixels are assigned to (Section 3.1) is decided using a distance measure  $D_s$  Equation (3).  $D_s$  is formed from the spatial distance  $d_s$  and the color distance  $d_c$  of the pixel and the cluster centers. The former measure Equation (1) is simple the Euclidean distance between the two coordinates, and the latter is the Euclidean distance between the lab colors of the two points Equation (2). We tested different color distance models to find the distance measure most suited for image fusion applications. In addition to the CIEDE2000 model (Section 3.1), the following color distance models are explored.

1. Euclidean color distance in the RGB color space: The simplest and the most well-known way to compute the distance between two colors is using the Euclidean distance in the RGB color space. The distance between two pixel colors  $C_1(R_1, G_1, B_1)$  and  $C_2(R_2, G_2, B_2)$  is computed as,

$$d = \sqrt{(R_2 - R_1)^2 + (G_2 - G_1)^2 + (B_2 - B_1)^2} \quad (7)$$

2. Color approximation distance (CAD): The perception of brightness by the human eye is non-linear, and for each color, non-linearity is not the same, as proven by different experiments [92–94]. Therefore, there have been many attempts to weight the RGB values to better fit human perception so that the color approximation would be more appropriate. The CAD between two colors  $C_1(R_1, G_1, B_1)$  and  $C_2(R_2, G_2, B_2)$  is computed as:

$$d = \sqrt{\left(2 + \frac{\bar{r}}{256}\right) \Delta R^2 + 4\Delta G^2 + \left(2 + \frac{255 - \bar{r}}{256}\right) \Delta B^2} \quad (8)$$

where  $\bar{r}$  is the average of  $R_1$  and  $R_2$ ;  $\Delta R$ ,  $\Delta G$ , and  $\Delta B$  are the differences between the red, green, and blue channels of the two colors, respectively.

3. CIEXYZ: The visualization of the RGB color space was not perfect because the color space of the human eye is greater than the experiment results of CIERGB, and an updated color difference model, CIEXYZ, was introduced by Commission Internationale de L'éclairage (CIE) [95]. It was mathematically designed to avoid the negative numbers known as tristimulus values. It consists of three parameters: non-negative cone response curve ( $X$ ), luminance ( $Y$ ), and partially-equal to blue color ( $Z$ ).

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.431 & 0.342 & 0.178 \\ 0.222 & 0.707 & 0.071 \\ 0.020 & 0.130 & 0.939 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The normalized tristimulus values  $x$ ,  $y$ , and  $z$  were calculated from  $X$ ,  $Y$ , and  $Z$  as:  $[x \ y \ z] = [\frac{X}{S} \ \frac{Y}{S} \ \frac{Z}{S}]$ , where  $S = X + Y + Z$ . The difference between the tristimulus values of two colors was then computed using the Euclidean distance formula.

4. CIE76: In the CIE76 color distance model, the CIELAB color space was used as it is considered to be the most exact means of representing color [96]. The color difference between two colors  $C_1(L_1, a_1, b_1)$  and  $C_2(L_2, a_2, b_2)$  in the CIELAB color space is calculated by following formula:

$$d = \sqrt{(L_1 - L_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2} \quad (9)$$

where  $\Delta L^* = L_1 - L_2$ ,  $\Delta a^* = a_1 - a_2$ , and  $\Delta b^* = b_1 - b_2$

5. CIE94: The CIE94 color model retains the CIELAB color space, and it addressed the non-uniformities in CIE76 [97,98]. CIE94 was defined in the lightness, chroma, and hue (LCh) color space, and they are calculated from the Lab-coordinates.
6. CMC l:c: The color difference CMC (Colour Measurement Committee of the Society of Dyes and Colourists of Great Britain) l:c is calculated in the LCh color space and has two parameters: lightness (l) and chroma (c). For the application, it allows the users based on the ratio of  $l : c$  to weight the difference that is deemed appropriate.
7. CIEDE2000: The CIE Delta E (CIEDE2000) color distance model was introduced by CIE and provides an improved method for the calculation of color differences. A detailed discussion is presented in Section 3.1.

The proposed algorithm with each of these distance measures was executed on the whole test dataset. The quality of the fused images was evaluated using all 12 fusion quality assessment metrics, and the average results are presented in Table 4. The best scores for each metric are highlighted in bold. The results show that most fusion quality assessment metrics rate the fusion results of the proposed method with the CIEDE2000 color distance model as the best. In these experiments, all other parameters were kept fixed. Moreover, we also computed the average execution time of the proposed method with each color distance model, and the results are reported in Table 4. One can note that CIEDE2000 takes the maximum time of 122.61 seconds per image pair due to its complex computations. However, it produces better decision maps for fusion and therefore is preferred over other measures.

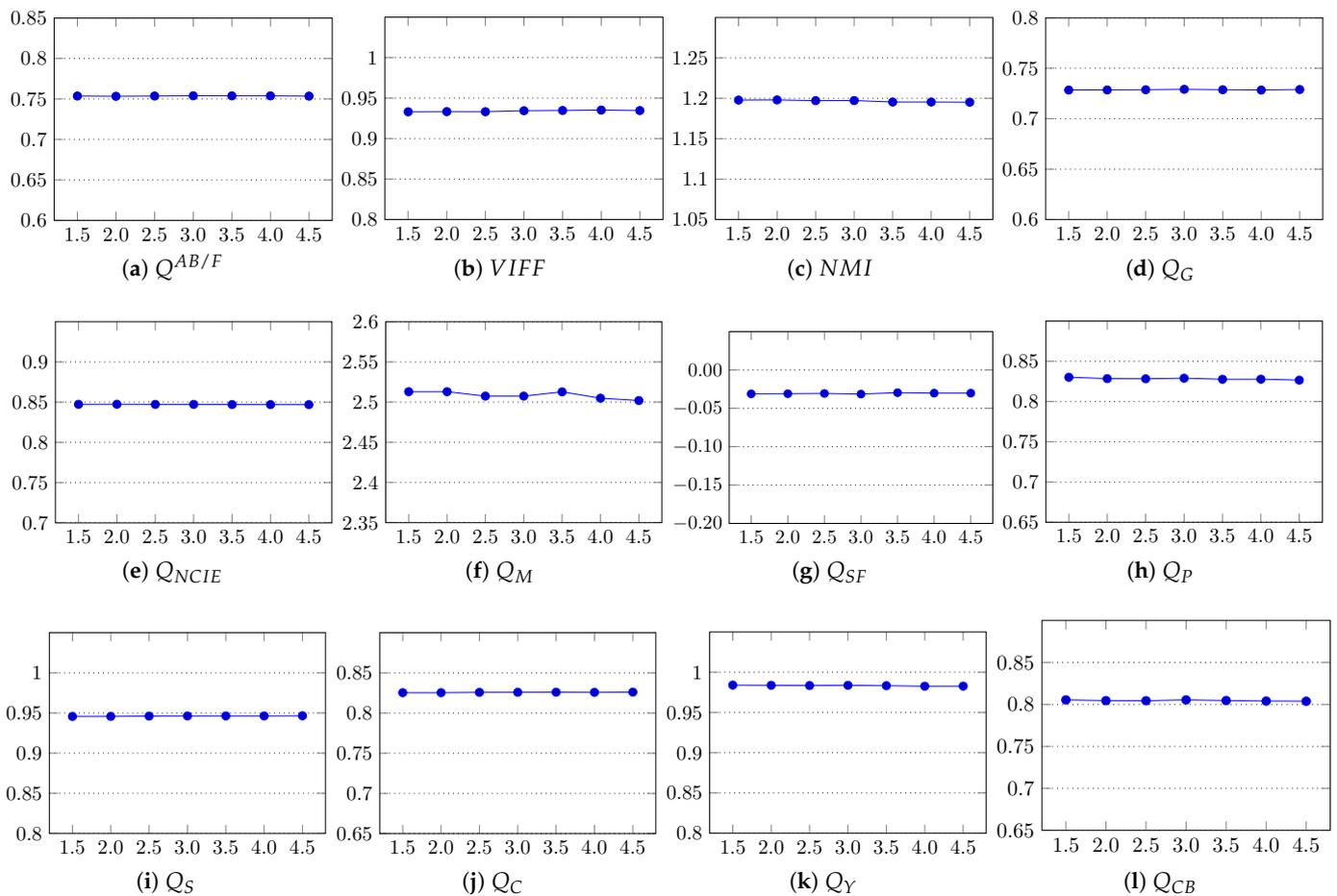
**Table 4.** Comparison of different color distance models for the proposed multi-focus image fusion algorithm. The best values are in bold. Time (in seconds) is the average execution time of the proposed method with each color distance model. CAD, color approximation distance.

Metric	Euclidean	CAD	CIE76	CIE94	CIEXYZ	CMC	CIEDE2000
$Q^{AB/F}$	0.7483	0.7477	0.7520	0.7510	0.7491	0.7533	<b>0.7539</b>
$VIFF$	0.9192	0.9182	0.9256	0.9240	0.9182	0.9287	<b>0.9344</b>
$NMI$	1.1919	1.1919	1.1945	1.1953	1.1953	1.1946	<b>1.1972</b>
$Q_G$	0.7252	0.7242	0.7291	0.7287	0.7287	<b>0.7295</b>	0.7291
$Q_{NCIE}$	0.8478	<b>0.8479</b>	0.8476	0.8478	<b>0.8479</b>	0.8472	0.8472
$Q_M$	2.4243	2.4406	2.4445	2.4363	2.3985	2.4603	<b>2.5074</b>
$Q_{SF}$	−0.0403	−0.0387	−0.0393	−0.0412	−0.0457	−0.0335	<b>−0.0314</b>
$Q_P$	0.8085	0.8077	0.8239	0.8239	0.8214	0.8224	<b>0.8289</b>
$Q_S$	0.9449	0.9449	0.9465	0.9463	0.9456	<b>0.9467</b>	0.9463
$Q_C$	0.8199	0.8190	0.8240	0.8238	0.8226	0.8250	<b>0.8258</b>
$Q_Y$	0.9803	0.9796	0.9827	0.9825	0.9822	0.9826	<b>0.9838</b>
$Q_{CB}$	0.7904	0.7881	0.8006	0.7994	0.7950	0.8005	<b>0.8054</b>
Time	52.08	56.87	44.71	45.31	<b>40.94</b>	53.67	122.61

## 5.2. Impact of the Number of Superpixels on the Fusion Quality

The parameter that is important in the computation of superpixels is  $k$  (Section 3.1), the desired number of approximately equally-sized superpixels. For an image with  $N$  pixels, the size of each superpixel is approximately  $N/k$  pixels. In this experiment, we analyzed the impact of the number of superpixels on the performance of the proposed algorithm.

We computed the fusion results with the proposed method using different values of  $k$ , and the results were evaluated using all 12 objective fusion quality metrics to get a general indicator. In particular, we tested the performance with  $k \in \{1000, 1500, 2000, 2500, 3000, 3500, 4000\}$ . From the results shown in Figure 10, one can note that most metrics show negligible or no change in the quality of the fused images; therefore, the impact of  $k$  on the fusion performance of our method is negligible. However, using a larger value of  $k$  certainly increases the execution time of the method due to the extra computation needed in computing the superpixels and their subsequent steps. The average execution times for different  $k$  are reported in Table 5. In our method,  $k = 3000$  was set as the default.



**Figure 10.** Impact of the number of superpixels  $k$  on the fusion quality of the proposed method. In all graphs, the x-axis represents  $k$  and the y-axis the quality score. The x-axis labels are coded as  $10^3$ .

**Table 5.** Execution time of the proposed method with different numbers of superpixels  $k$ . Time (in seconds) is the average execution time. The best time is marked in bold.

$k$	1500	2000	2500	3000	3500	4000	4500
Time	<b>97.86</b>	109.16	118.69	122.61	119.98	132.39	131.43

### 5.3. Analysis of the Relative Contribution of Spatial and Color Distances

The parameter  $m$  in Equation (3) is used to control the compactness of the superpixels: it weights the relative information between the spatial and the color proximity. The recommended range of the value of  $m$  in the CIELAB color space is  $[1, 40]$ . We performed a set of experiments using different values of weighting factor  $m$ . We note that for small values of  $m$ , the shape and size of the superpixels become less regular and more tightly adhere to the image boundaries. For a large value of  $m$ , the superpixels are more compact, and spatial proximity becomes more important. The results of superpixel estimations with different values of  $m$  on a sample image are shown in Figure 11.



Figure 11. Superpixels estimated with the proposed algorithm using (a)  $m = 15$  and (b)  $m = 40$ .

The fusion results were assessed using all 12 quality metrics, and their average scores are graphed in Figure 12. The results clearly show that  $m$  does not contribute to the quality enhancement of the fused image. Since for large values of  $m$ , the superpixels are more compact, the execution time of the method is reduced, as shown in Table 6. In the implementation of our method, the default value of  $m$  was set to 25.

Table 6. Impact of weight factor  $m$  on execution time. Time (in seconds) is the average execution time of the proposed method with different values of  $m$ . The best values are in bold.

$m$	10	15	20	25	30	35	40
Time	151.72	141.16	125.76	122.61	110.32	110.32	<b>98.54</b>

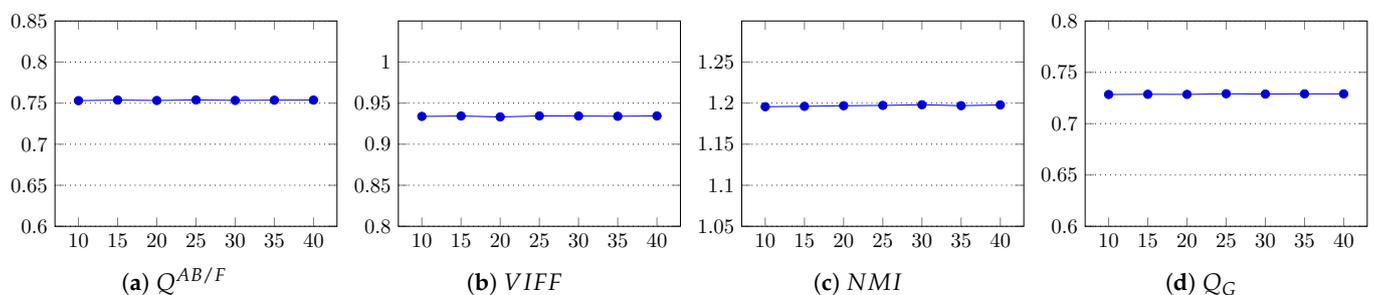
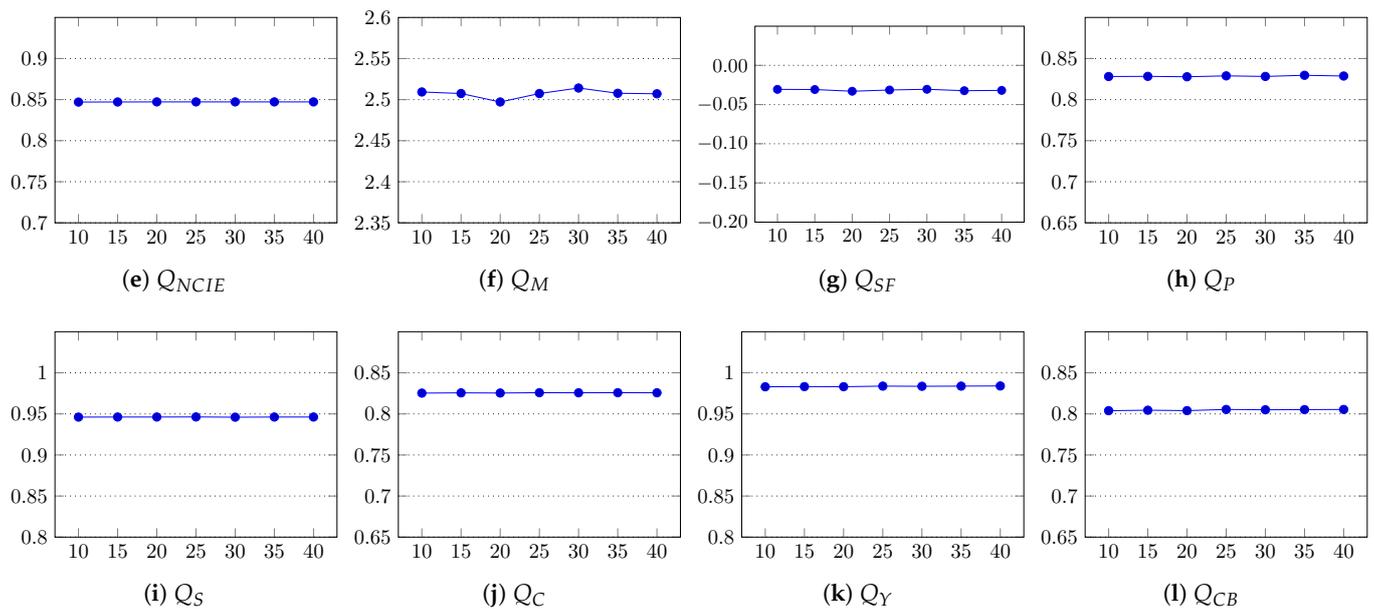


Figure 12. Cont.



**Figure 12.** Impact of the weight factor  $m$  on the fusion quality of the proposed method. In all graphs, the x-axis represents  $m$  and the y-axis the quality score.

## 6. Conclusions

In this paper, a novel multi-focus image fusion technique is presented. The proposed technique merges similar pixels into larger groups termed superpixels. The spatial and color properties of the pixels are exploited to estimate the superpixels in the source multi-focus images. Moreover, different color distance models are tested to obtain high quality fused images. Different statistical properties of the computed superpixels are analyzed to categorize them as focused and de-focused, and based on this information, a focus map is generated. Moreover, the wrongly categorized superpixels are corrected by a spatial constraint rule. The qualitative and quantitative experiments worked out on a benchmark multi-focus image dataset reveal that our method produces better quality images than existing similar techniques. The source code of the proposed method is released publicly at the project website (<http://www.di.unito.it/~farid/Research/SBF.html>, accessed on 18 February 2021).

**Author Contributions:** Conceptualization, A.I., M.S.F., M.H.K. and M.G.; methodology, A.I., M.S.F. and M.H.K.; software, A.I., M.S.F. and M.H.K.; validation, A.I., M.S.F., M.H.K. and M.G.; investigation, A.I., M.S.F., M.H.K.; writing—original draft preparation, A.I., M.S.F., M.H.K. and M.G.; writing—review and editing, A.I., M.S.F., M.H.K., and M.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data sharing not applicable.

**Conflicts of Interest:** The authors declare that they have no conflict of interest.

## References

1. Wan, T.; Zhu, C.; Qin, Z. Multifocus image fusion based on robust principal component analysis. *Pattern Recognit. Lett.* **2013**, *34*, 1001–1008. [[CrossRef](#)]
2. Campbell, J.B.; Wynne, R.H. *Introduction to Remote Sensing*; Guilford Press: New York, NY, USA, 2011.

3. Ardeshir Goshtasby, A.; Nikolov, S. Guest editorial: Image fusion: Advances in the state of the art. *Inf. Fusion* **2007**, *8*, 114–118. [[CrossRef](#)]
4. Wang, X.; Bai, S.; Li, Z.; Sui, Y.; Tao, J. The PAN and MS image fusion algorithm based on adaptive guided filtering and gradient information regulation. *Inf. Sci.* **2021**, *545*, 381–402. [[CrossRef](#)]
5. Xu, Y.; Yang, C.; Sun, B.; Yan, X.; Chen, M. A novel multi-scale fusion framework for detail-preserving low-light image enhancement. *Inf. Sci.* **2021**, *548*, 378–397. [[CrossRef](#)]
6. Du, J.; Li, W.; Tan, H. Three-layer medical image fusion with tensor based features. *Inf. Sci.* **2020**, *525*, 93–108. [[CrossRef](#)]
7. Li, S.; Yang, B.; Hu, J. Performance comparison of different multi-resolution transforms for image fusion. *Inf. Fusion* **2011**, *12*, 74–84. [[CrossRef](#)]
8. Zafar, R.; Farid, M.S.; Khan, M.H. Multi-Focus Image Fusion: Algorithms, Evaluation, and a Library. *J. Imaging* **2020**, *6*, 60. [[CrossRef](#)]
9. Saha, A.; Bhatnagar, G.; Wu, Q.J. Mutual spectral residual approach for multifocus image fusion. *Digit. Signal Process.* **2013**, *23*, 1121–1135. [[CrossRef](#)]
10. Goshtasby, A.A.; Nikolov, S. Image fusion: advances in the state of the art. *Inf. Fusion* **2007**, *2*, 114–118. [[CrossRef](#)]
11. Li, S.; Kwok, J.T.; Wang, Y. Multifocus image fusion using artificial neural networks. *Pattern Recognit. Lett.* **2002**, *23*, 985–997. [[CrossRef](#)]
12. Eltoukhy, H.A.; Kavusi, S. Computationally efficient algorithm for multifocus image reconstruction. In Proceedings of the Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications IV, Santa Clara, CA, USA, 20–24 January 2003; Volume 5017, pp. 332–341.
13. Liu, Y.; Liu, S.; Wang, Z. Multi-focus image fusion with dense SIFT. *Inf. Fusion* **2015**, *23*, 139–155. [[CrossRef](#)]
14. Li, S.; Kang, X.; Fang, L.; Hu, J.; Yin, H. Pixel-level image fusion: A survey of the state of the art. *Inf. Fusion* **2017**, *33*, 100–112. [[CrossRef](#)]
15. Li, S.; Kwok, J.Y.; Tsang, I.H.; Wang, Y. Fusing images with different focuses using support vector machines. *IEEE Trans. Neural Netw.* **2004**, *15*, 1555–1561. [[CrossRef](#)]
16. Huang, W.; Jing, Z. Multi-focus image fusion using pulse coupled neural network. *Pattern Recognit. Lett.* **2007**, *28*, 1123–1132. [[CrossRef](#)]
17. Agrawal, D.; Singhai, J. Multifocus image fusion using modified pulse coupled neural network for improved image quality. *IET Image Process.* **2010**, *4*, 443–451. [[CrossRef](#)]
18. Li, S.; Kang, X.; Hu, J.; Yang, B. Image matting for fusion of multi-focus images in dynamic scenes. *Inf. Fusion* **2013**, *14*, 147–162. [[CrossRef](#)]
19. Kumar, B.S. Image fusion based on pixel significance using cross bilateral filter. *Signal Image Video Process.* **2015**, *9*, 1193–1204. [[CrossRef](#)]
20. Liu, Y.; Wang, L.; Cheng, J.; Li, C.; Chen, X. Multi-focus image fusion: A Survey of the state of the art. *Inf. Sci.* **2020**, *64*, 71–91. [[CrossRef](#)]
21. Piella, G. A general framework for multiresolution image fusion: From pixels to regions. *Inf. Fusion* **2003**, *4*, 259–280. [[CrossRef](#)]
22. Petrovic, V.; Xydeas, C.S. Gradient based multiresolution image fusion. *IEEE Trans. Image Process.* **2004**, *13*, 228–237. [[CrossRef](#)]
23. Li, H.; Manjunath, B.; Mitra, S.K. Multisensor image fusion using the wavelet transform. *Graph Model Image Proc.* **1995**, *57*, 235–245. [[CrossRef](#)]
24. Kumar, B.S. Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform. *Signal Image Video Process.* **2013**, *7*, 1125–1143. [[CrossRef](#)]
25. Nencini, F.; Garzelli, A.; Baronti, S.; Alparone, L. Remote sensing image fusion using the curvelet transform. *Inf. Fusion* **2007**, *8*, 143–156. [[CrossRef](#)]
26. Xiao-Bo, Q.; Jing-Wen, Y.; Hong-Zhi, X.; Zi-Qian, Z. Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampling contourlet transform domain. *Acta Autom. Sin.* **2008**, *34*, 1508–1514.
27. Zhang, Q.; Guo, B.L. Multifocus image fusion using the nonsubsampling contourlet transform. *Signal Process.* **2009**, *89*, 1334–1346. [[CrossRef](#)]
28. Bavirisetti, D.P.; Dhuli, R. Multi-focus image fusion using multi-scale image decomposition and saliency detection. *Ain Shams Eng. J.* **2018**, *9*, 1103–1117. [[CrossRef](#)]
29. Zheng, S.; Shi, W.Z.; Liu, J.; Zhu, G.X.; Tian, J.W. Multisource image fusion method using support value transform. *IEEE Trans. Image Process.* **2007**, *16*, 1831–1839. [[CrossRef](#)]
30. Haghghat, M.B.A.; Aghagolzadeh, A.; Seyedarabi, H. Multi-focus image fusion for visual sensor networks in DCT domain. *Comput. Electr. Eng.* **2011**, *37*, 789–797. [[CrossRef](#)]
31. Naidu, V.; Elias, B. A novel image fusion technique using DCT based Laplacian pyramid. *Int. J. Invent. Eng. Sci.* **2013**, *5*, 2319–2398.
32. Bai, X.; Liu, M.; Chen, Z.; Wang, P.; Zhang, Y. Morphology and active contour model for multi-focus image fusion. In Proceedings of the International Conference on Digital Image Computing: Techniques and Applications (DICTA), Adelaide, SA, Australia, 23–25 November 2015; pp. 1–6.
33. Yang, B.; Li, S. Multifocus image fusion and restoration with sparse representation. *IEEE Trans. Instrum. Meas.* **2009**, *59*, 884–892. [[CrossRef](#)]

34. Liang, J.; He, Y.; Liu, D.; Zeng, X. Image fusion using higher order singular value decomposition. *IEEE Trans. Image Process.* **2012**, *21*, 2898–2909. [[CrossRef](#)]
35. Luo, X.; Zhang, Z.; Zhang, C.; Wu, X. Multi-focus image fusion using HOSVD and edge intensity. *J. Vis. Commun. Image Represent.* **2017**, *45*, 46–61. [[CrossRef](#)]
36. Zhang, X.; Li, X.; Liu, Z.; Feng, Y. Multi-focus image fusion using image-partition based focus detection. *Signal Process.* **2014**, *102*, 64–76. [[CrossRef](#)]
37. Li, S.; Kwok, J.T.; Wang, Y. Combination of images with diverse focuses using the spatial frequency. *Inf. Fusion* **2001**, *2*, 169–176. [[CrossRef](#)]
38. Aslantas, V.; Kurban, R. Fusion of multi-focus images using differential evolution algorithm. *Expert Syst. Appl.* **2010**, *37*, 8861–8870. [[CrossRef](#)]
39. Siddique, A.; Xiao, B.; Li, W.; Nawaz, Q.; Hamid, I. Multi-Focus Image Fusion Using Block-Wise Color-Principal Component Analysis. In Proceedings of the IEEE 3rd International Conference on Image, Vision and Computing (ICIVC), Chongqing, China, 27–29 June 2018; pp. 458–462.
40. Li, M.; Cai, W.; Tan, Z. A region based multi-sensor image fusion scheme using pulse-coupled neural network. *Pattern Recognit. Lett.* **2006**, *27*, 1948–1956. [[CrossRef](#)]
41. Li, S.; Yang, B. Multifocus image fusion using region segmentation and spatial frequency. *Image Vis. Comput.* **2008**, *26*, 971–979. [[CrossRef](#)]
42. Lewis, J.J.; O’Callaghan, R.J.; Nikolov, S.G.; Bull, D.R.; Canagarajah, N. Pixel-and region based image fusion with complex wavelets. *Inf. Fusion* **2007**, *8*, 119–130. [[CrossRef](#)]
43. Nejati, M.; Samavi, S.; Shirani, S. Multi-focus image fusion using dictionary based sparse representation. *Inf. Fusion* **2015**, *25*, 72–84. [[CrossRef](#)]
44. Huang, W.; Jing, Z. Evaluation of focus measures in multi-focus image fusion. *Pattern Recognit. Lett.* **2007**, *28*, 493–500. [[CrossRef](#)]
45. Zhou, Z.; Li, S.; Wang, B. Multi-scale weighted gradient based fusion for multi-focus images. *Inf. Fusion* **2014**, *20*, 60–72. [[CrossRef](#)]
46. Li, S.; Kang, X.; Hu, J. Image fusion with guided filtering. *IEEE Trans. Image Process.* **2013**, *22*, 2864–2875.
47. Bouzos, O.; Andreadis, I.; Mitianoudis, N. Conditional random field model for robust multi-focus image fusion. *IEEE Trans. Image Process.* **2019**, *28*, 5636–5648. [[CrossRef](#)] [[PubMed](#)]
48. Chen, Y.; Guan, J.; Cham, W.K. Robust multi-focus image fusion using edge model and multi-matting. *IEEE Trans. Image Process.* **2017**, *27*, 1526–1541. [[CrossRef](#)]
49. Hua, K.L.; Wang, H.C.; Rusdi, A.H.; Jiang, S.Y. A novel multi-focus image fusion algorithm based on random walks. *J. Vis. Commun. Image Represent.* **2014**, *25*, 951–962. [[CrossRef](#)]
50. Li, H.; Nie, R.; Cao, J.; Guo, X.; Zhou, D.; He, K. Multi-Focus Image Fusion Using U-Shaped Networks With a Hybrid Objective. *IEEE Sens. J.* **2019**, *19*, 9755–9765. [[CrossRef](#)]
51. Liu, Y.; Chen, X.; Wang, Z.; Wang, Z.J.; Ward, R.K.; Wang, X. Deep learning for pixel-level image fusion: Recent advances and future prospects. *Inf. Fusion* **2018**, *42*, 158–173. [[CrossRef](#)]
52. Tang, H.; Xiao, B.; Li, W.; Wang, G. Pixel convolutional neural network for multi-focus image fusion. *Inf. Sci.* **2018**, *433–434*, 125–141. [[CrossRef](#)]
53. Mustafa, H.T.; Yang, J.; Zareapoor, M. Multi-scale convolutional neural network for multi-focus image fusion. *Image Vis. Comput.* **2019**, *85*, 26–35. [[CrossRef](#)]
54. Liu, Y.; Chen, X.; Peng, H.; Wang, Z. Multi-focus image fusion with a deep convolutional neural network. *Inf. Fusion* **2017**, *36*, 191–207. [[CrossRef](#)]
55. Mustafa, H.T.; Zareapoor, M.; Yang, J. MLDNet: Multi-level dense network for multi-focus image fusion. *Signal Process. Image Commun.* **2020**, *85*, 115864. [[CrossRef](#)]
56. Wang, C.; Zhao, Z.; Ren, Q.; Xu, Y.; Yu, Y. A novel multi-focus image fusion by combining simplified very deep convolutional networks and patch based sequential reconstruction strategy. *Appl. Soft Comput.* **2020**, *91*, 106253. [[CrossRef](#)]
57. Liu, Y.; Chen, X.; Ward, R.K.; Wang, Z.J. Medical Image Fusion via Convolutional Sparsity Based Morphological Component Analysis. *IEEE Signal Process. Lett.* **2019**, *26*, 485–489. [[CrossRef](#)]
58. Zhang, C. Multi-focus Image Fusion Based on Convolutional Sparse Representation with Mask Simulation. In *Advances in 3D Image and Graphics Representation, Analysis, Computing and Information Technology*; Kountchev, R., Patnaik, S., Shi, J., Favorskaya, M.N., Eds.; Springer: Singapore, 2020; pp. 159–168.
59. Xing, C.; Wang, M.; Dong, C.; Duan, C.; Wang, Z. Using Taylor Expansion and Convolutional Sparse Representation for Image Fusion. *Neurocomputing* **2020**, *402*, 437–455. [[CrossRef](#)]
60. Zhang, H.; Le, Z.; Shao, Z.; Xu, H.; Ma, J. MFF-GAN: An unsupervised generative adversarial network with adaptive and gradient joint constraints for multi-focus image fusion. *Inf. Fusion* **2021**, *66*, 40–53. [[CrossRef](#)]
61. Huang, J.; Le, Z.; Ma, Y.; Mei, X.; Fan, F. A generative adversarial network with adaptive constraints for multi-focus image fusion. *Neural Comput. Appl.* **2020**, *32*, 15119–15129. [[CrossRef](#)]
62. Murugan, A.; Arumugam, G.; Gobinath, D. Multi-Focus Image Fusion Using Conditional Generative Adversarial Networks. In *Intelligent Computing and Applications*; Dash, S.S., Das, S., Panigrahi, B.K., Eds.; Springer: Singapore, 2021; pp. 559–566.

63. Veksler, O.; Boykov, Y.; Mehrani, P. Superpixels and Supervoxels in an Energy Optimization Framework. In *Computer Vision—ECCV 2010, Proceedings of the European Conference on Computer Vision (ECCV 2010), Heraklion, Crete, Greece, 5–11 September 2010*; Daniilidis, K., Maragos, P., Paragios, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 211–224.
64. Levinshtein, A.; Stere, A.; Kutulakos, K.N.; Fleet, D.J.; Dickinson, S.J.; Siddiqi, K. TurboPixels: Fast Superpixels Using Geometric Flows. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 2290–2297. [[CrossRef](#)]
65. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2274–2282. [[CrossRef](#)] [[PubMed](#)]
66. Sharma, G.; Wu, W.; Dalal, E.N. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Res. Appl.* **2005**, *30*, 21–30. [[CrossRef](#)]
67. Kumar, V.; Gupta, P. Importance of statistical measures in digital image processing. *Int. J. Emerg. Technol. Adv. Eng.* **2012**, *2*, 56–62.
68. Wang, W.; Chang, F. A Multi-focus Image Fusion Method Based on Laplacian Pyramid. *J. Comput.* **2011**, *6*, 2559–2566. [[CrossRef](#)]
69. Farid, M.S.; Mahmood, A.; Al-Maadeed, S.A. Multi-focus image fusion using Content Adaptive Blurring. *Inf. Fusion* **2019**, *45*, 96–112. [[CrossRef](#)]
70. Lytro Multi-Focus Dataset. Available online: <http://mansournejati.ece.iut.ac.ir/content/lytro-multi-focus-dataset> (accessed on 31 October 2019).
71. Tian, J.; Chen, L. Adaptive multi-focus image fusion using a wavelet based statistical sharpness measure. *Signal Process.* **2012**, *92*, 2137–2146. [[CrossRef](#)]
72. Qu, X.; Hu, C.; Yan, J. Image fusion algorithm based on orientation information motivated pulse coupled neural networks. In *Proceedings of the 7th World Congress on Intelligent Control and Automation, Chongqing, China, 25–27 June 2008*; pp. 2437–2441.
73. Mitianoudis, N.; Stathaki, T. Pixel-based and region-based image fusion schemes using ICA bases. *Inf. Fusion* **2007**, *8*, 131–142. [[CrossRef](#)]
74. Paul, S.; Sevcenco, I.S.; Agathoklis, P. Multi-exposure and multi-focus image fusion in gradient domain. *J. Circuits Syst. Comput.* **2016**, *25*, 1650123. [[CrossRef](#)]
75. Naidu, V.; Raol, J.R. Pixel-level image fusion using wavelets and principal component analysis. *Def. Sci. J.* **2008**, *58*, 338–352. [[CrossRef](#)]
76. Liu, Y.; Chen, X.; Ward, R.K.; Wang, Z.J. Image fusion with convolutional sparse representation. *IEEE Signal Process. Lett.* **2016**, *23*, 1882–1886. [[CrossRef](#)]
77. Liu, Z.; Blasch, E.; Xue, Z.; Zhao, J.; Laganieri, R.; Wu, W. Objective Assessment of Multiresolution Image Fusion Algorithms for Context Enhancement in Night Vision: A Comparative Study. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 94–109. [[CrossRef](#)]
78. Qu, G.; Zhang, D.; Yan, P. Information measure for performance of image fusion. *Electron. Lett.* **2002**, *38*, 313–315. [[CrossRef](#)]
79. Hossny, M.; Nahavandi, S.; Creighton, D. Comments on ‘Information measure for performance of image fusion’. *Electron. Lett.* **2008**, *44*, 1066–1067. [[CrossRef](#)]
80. Wang, Q.; Shen, Y.; Jin, J. 19—Performance evaluation of image fusion techniques. In *Image Fusion*; Stathaki, T., Ed.; Academic Press: Oxford, UK, 2008; pp. 469–492.
81. Han, Y.; Cai, Y.; Cao, Y.; Xu, X. A new image fusion performance metric based on visual information fidelity. *Inf. Fusion* **2013**, *14*, 127–135. [[CrossRef](#)]
82. Xydeas, C.S.; Petrovic, V. Objective image fusion performance measure. *Electron. Lett.* **2000**, *36*, 308–309. [[CrossRef](#)]
83. Xydeas, C.S.; Petrovic, V. Objective pixel-level image fusion performance measure. *Int. Soc. Opt. Photonics* **2000**, *4051*, pp. 89–98.
84. Zheng, Y.; Essock, E.A.; Hansen, B.C.; Haun, A.M. A new metric based on extended spatial frequency and its application to DWT based fusion algorithms. *Inf. Fusion* **2007**, *8*, 177–192. [[CrossRef](#)]
85. Wang, P.W.; Liu, B. A novel image fusion metric based on multi-scale analysis. In *Proceedings of the 2008 9th International Conference on Signal Processing, Beijing, China, 26–29 October 2008*; pp. 965–968.
86. Liu, Z.; Forsyth, D.S.; Laganieri, R. A feature based metric for the quantitative evaluation of pixel-level image fusion. *Comput. Vis. Image Underst.* **2008**, *109*, 56–68. [[CrossRef](#)]
87. Piella, G.; Heijmans, H. A new quality metric for image fusion. In *Proceedings of the 2003 International Conference on Image Processing (Cat. No. 03CH37429), Barcelona, Spain, 14–17 September 2003*.
88. Cvejic, N.; Loza, A.; Bull, D.; Canagarajah, N. A similarity metric for assessment of image fusion algorithms. *Int. J. Signal Process.* **2005**, *2*, 178–182.
89. Yang, C.; Zhang, J.Q.; Wang, X.R.; Liu, X. A novel similarity based quality metric for image fusion. *Inf. Fusion* **2008**, *9*, 156–160. [[CrossRef](#)]
90. Chen, Y.; Blum, R.S. A new automated quality assessment algorithm for image fusion. *Image Vis. Comput.* **2009**, *27*, 1421–1432. [[CrossRef](#)]
91. Zhao, J.; Laganieri, R.; Liu, Z. Performance assessment of combinative pixel-level image fusion based on an absolute feature measurement. *Int. J. Innov. Comput. Inf. Control.* **2007**, *3*, 1433–1447.
92. Industrial Colour-Difference Evaluation. Available online: <http://cie.co.at/publications/industrial-colour-difference-evaluation> (accessed on 1 January 2021).
93. Colour Metric. Available online: [www.compuphase.com/cmtric.htm](http://www.compuphase.com/cmtric.htm) (accessed on 15 October 2020).
94. Kuehni, R.G. Industrial color difference: Progress and problems. *Color Res. Appl.* **1990**, *15*, 261–265. [[CrossRef](#)]
95. Smith, T.; Guild, J. The C.I.E. colorimetric standards and their use. *Trans. Opt. Soc.* **1931**, *33*, 73–134. [[CrossRef](#)]

- 
96. Sharma, G.; Bala, R. *Digital Color Imaging Handbook*; CRC press: Boca Raton, FL, USA, 2017.
  97. Melgosa, M. CIE94, History, Use, and Performance. In *Encyclopedia of Color Science and Technology*; Springer: New York, NY, USA, 2014; pp. 1–5.
  98. McDonald, R.; Smith, K.J. CIE94—a new colour-difference formula. *J. Soc. Dye. Colour.* **1995**, *111*, 376–379. [[CrossRef](#)]