

Communication

Multi-Focus Image Fusion Using Focal Area Extraction in a Large Quantity of Microscopic Images

Jiyoung Lee ^{1,†}, Seunghyun Jang ^{1,†}, Jungbin Lee ², Taehan Kim ¹, Seonghan Kim ², Jongbum Seo ¹, Ki Hean Kim ^{2,*} and Sejung Yang ^{1,*}

¹ Department of Biomedical Engineering, College of Software and Digital Healthcare Convergence, Yonsei University, Wonju 26493, Korea; amyjylee95@yonsei.ac.kr (J.L.); jangbi3486@yonsei.ac.kr (S.J.); taehan5479@gmail.com (T.K.); jongbums@yonsei.ac.kr (J.S.)

² Department of Mechanical Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Korea; jungbinlee@postech.ac.kr (J.L.); seonghan@postech.ac.kr (S.K.)

* Correspondence: kiheankim@postech.ac.kr (K.H.K.); syang@yonsei.ac.kr (S.Y.)

† These authors contributed equally to this work.

Abstract: The non-invasive examination of conjunctival goblet cells using a microscope is a novel procedure for the diagnosis of ocular surface diseases. However, it is difficult to generate an all-in-focus image due to the curvature of the eyes and the limited focal depth of the microscope. The microscope acquires multiple images with the axial translation of focus, and the image stack must be processed. Thus, we propose a multi-focus image fusion method to generate an all-in-focus image from multiple microscopic images. First, a bandpass filter is applied to the source images and the focus areas are extracted using Laplacian transformation and thresholding with a morphological operation. Next, a self-adjusting guided filter is applied for the natural connections between local focus images. A window-size-updating method is adopted in the guided filter to reduce the number of parameters. This paper presents a novel algorithm that can operate for a large quantity of images (10 or more) and obtain an all-in-focus image. To quantitatively evaluate the proposed method, two different types of evaluation metrics are used: “full-reference” and “no-reference”. The experimental results demonstrate that this algorithm is robust to noise and capable of preserving local focus information through focal area extraction. Additionally, the proposed method outperforms state-of-the-art approaches in terms of both visual effects and image quality assessments.

Keywords: image fusion; all-in-focus; depth of field; microscopy



Citation: Lee, J.; Jang, S.; Lee, J.; Kim, T.; Kim, S.; Seo, J.; Kim, K.H.; Yang, S. Multi-Focus Image Fusion Using Focal Area Extraction in a Large Quantity of Microscopic Images. *Sensors* **2021**, *21*, 7371. <https://doi.org/10.3390/s21217371>

Academic Editors: Christoph M. Friedrich, Filippo Piccinini, Antonella Carbonaro, Peter Horvath and Gastone C. Castellani

Received: 25 September 2021

Accepted: 3 November 2021

Published: 5 November 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

Generating all-in-focus images is the process of combining visual information from multiple input images into a single image. The resulting image must contain more accurate, stable, and complete information than the input images, and N sets of sub-images from different in-focus images are used to obtain the resulting images, from which all focus areas are fused [1]. This process is accomplished by using multi-focus image fusion (MFIF) techniques and is observed in various fields, including digital photography and medical diagnosis [2].

The non-invasive examination of the conjunctiva using a microscope is a state-of-the-art method to diagnose ocular surface diseases. It is performed by observing and analyzing conjunctival goblet cells, which secrete mucins on the ocular surface to form the mucus layer of the tear film. The mucus layer is important for tear film stability, and many ocular surface diseases are associated with tear film instability. In confocal microscopy, the axial resolution often misses important information in areas when the subject is out of focus, owing to a shallow depth of field (DOF) and small field of view (FOV) up to $500\ \mu\text{m} \times 500\ \mu\text{m}$ [3,4]. Confocal microscopy includes limitations, such as a relatively slow imaging speed due to the point-scanning method. A wide-field fluorescence microscopy that improves the

existing limitations was developed for the non-invasive imaging of conjunctival goblet cells [5]. The new fluorescence microscopy visualizes conjunctival goblet cells in high contrasts via fluorescence labeling with moxifloxacin antibiotic ophthalmic solution. It is specialized for live animal models based on its fast imaging speed and large FOV of 1.6 mm × 1.6 mm, and it has the potential for clinical applications. Nevertheless, a high DOF was required to examine the goblet cells in the tilted conjunctiva. Even the most focused images contain unfocused areas, which implies that they lack important information. To solve this problem, it is necessary to obtain several local focus images with different focus areas, and to then combine them into all-in-focus images.

The MFIF method is mainly divided into the transform-domain and spatial-domain methods [6,7]. Transform-domain methods include image transformation, coefficient fusion, and inverse transformation. Source images are converted into a transform domain, and then the transformed coefficients are merged using a fusion strategy. Li et al. introduced the discrete wavelet transform (DWT) into image fusion [8]. The DWT image fusion method consists of three stages: wavelet transformation, maximum selection, and image fusion. Their method fuses wavelet coefficients using maximum selection based on the absolute values of the maximum values in each window. The values of the wavelet coefficients are then adjusted using a filter, according to the ambient values. However, the DWT does not satisfy shift invariance, which is one of the most important characteristics of image fusion, resulting in incorrect fusion or noise. To solve this problem, an image fusion technique based on the shift-invariant DWT model was proposed, and it achieved better results than the original DWT-based method [9]. In addition to the image fusion methods discussed above, the image fusion techniques using transform-domain methods, such as independent component analyses, discrete cosine transformation, and hybrid image fusion methods combining wavelet transformation and curve transformation were also proposed [10,11].

In spatial-domain methods, source images are fused based on the spatial features of the images. Images are mainly fused using pixel values; such methods are simple to implement and can preserve large amounts of information [12]. Li et al. also introduced a spatial-domain image fusion method based on block division [13]. In this method, the input images are divided into several blocks of a fixed size, and threshold-based fusion rules are applied to obtain the fused blocks. Block-based methods can be enhanced by including threshold processing and block segmentation. Block-based image fusion methods fix the block size that affects the fusion results. To solve this problem, adaptive block segmentation methods with different block sizes can be implemented for each input image. The adaptive block method is a quad-tree, block-based method [14]. This method decomposes input images into a quad-tree structure and then detects the focal areas within each block. Additionally, a region-based image fusion method was developed to increase the flexibility of input images. This method subdivides input images into super pixels using both block-based and region-based characteristics simultaneously [15,16]. The basic goal of image fusion is to improve the visual quality of fused images by dividing the boundaries between focused and defocused areas in the input images. In addition to the transform-domain and spatial-domain methods, various hybrid methods and deep learning methods were proposed [17–19].

In this paper, we propose a novel MFIF method that analyzes sequences of up to 20 microscopy input images corresponding to different DOF levels. This method is optimized for the newly developed microscope and can analyze goblet cells through results with high DOF. We solve the problems in both the transform domain and spatial domain and present a method for image fusion based on focus area detection. To evaluate the effectiveness of the proposed method, we conduct the application of our method to camera images and conjunctival goblet cell images.

2. Materials and Methods

2.1. Proposed Method

Figure 1 presents a schematic diagram of an image fusion method including focal area extraction. The proposed method is applicable to a large quantity of local-focus images to generate an all-in-focus image. Let I_n be the set of input image sets. First, we adopt a band-pass filter to all filters of the input image sets to enhance the gradient information and edges of the local-focus areas. We then utilize Laplacian filters to enhance the focus areas and thresholding to extract the focus areas, which are denoted as I_{thn} . Next, a guided filter is applied, after removing unnecessary areas, by dilating the focus areas. Finally, the focus areas, I_{gn} , outputted by the guided filter are combined using the pixel-wise weighted averaging rule and an all-in-focus image is obtained.

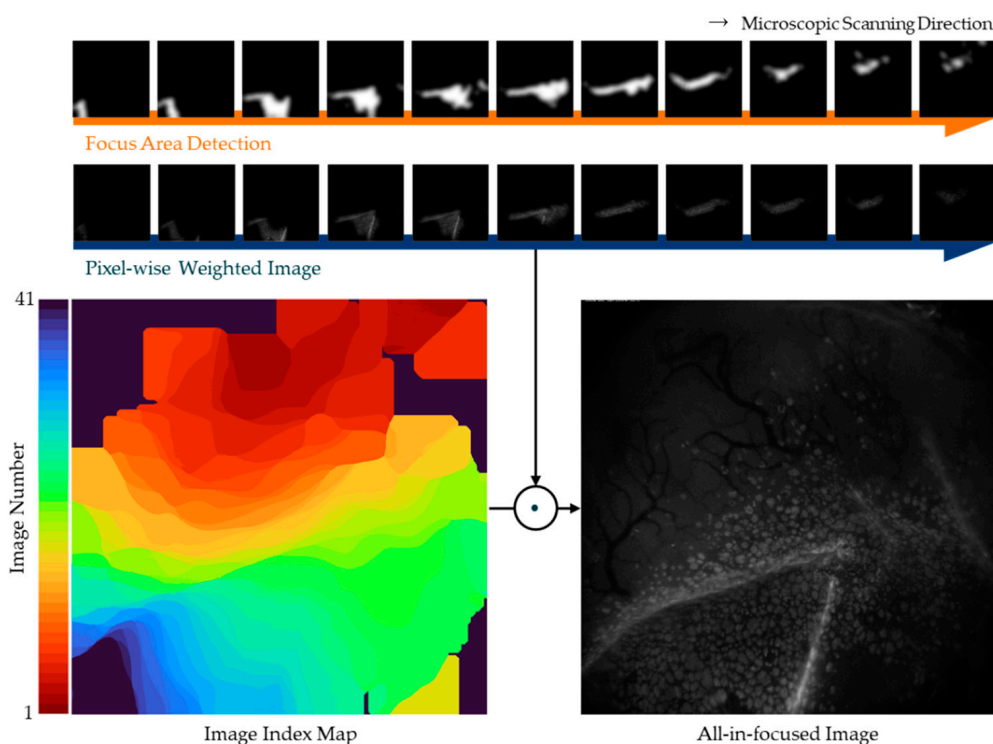


Figure 1. Schematic diagram of the proposed method.

2.2. Subjects

In this study, moxifloxacin-based, axially swept, wide-field fluorescence microscopy (WFFM) was employed. The objective lens was initially positioned so that the focal plane was at the deepest location of the specimen surface. Then, the focal plane was swept outward by the translation of the objective lens with continuous WFFM imaging. The imaging field of view (FOV) was $1.6 \text{ mm} \times 1.6 \text{ mm}$, the image resolution was $1.3 \mu\text{m}$, and imaging speed was 30 frames/s. The WFFM system had a shallow DOF of approximately $30 \mu\text{m}$. Typical images had 2048×2048 gray scale pixels. Seven 8-week-old SKH1-Hrhr male mice were used for in vivo GC imaging experiment [5].

2.3. Focus Area Enhancement Based on the Transform Domain

Local focus images obtained using a microscope require denoising and focus area extraction. A defocus area has a narrower bandwidth than a focus area [20]. Therefore, a focus area has higher frequency information than a defocus area [21,22]. This section presents a method for enhancing focus areas handling the transformation domain.

For domain transformation, we used a Fourier transform to extract high-frequency information and perform denoising simultaneously by applying a band-pass filter. This filter was designed in a Gaussian form, and an appropriate cutoff frequency value was set:

$$F_n(u, v) = fft(I_n(x, y)), n = 1 \dots N, \quad (1)$$

$$I_{bpn}(x, y) = fft^{-1}(H_b(u, v)F_n(u, v)), n = 1 \dots N, \quad (2)$$

where I_n is an input image with N datasets, and fft denotes a Fourier transform; I_{bpn} is a result image with denoising and focus area enhancement performed using the band-pass filter.

2.4. Focus Area Detection

After deriving I_{bpn} using a band-pass filter, we applied a Laplacian filter, which is an edge detection method. This filter was employed to compute the second derivative of an image by measuring the rate at which the first derivative changes. This determined whether a change in adjacent pixel values was caused by an edge or continuous progression [23]:

$$I_{ln}(x, y) = L(I_{bpn}(x, y), M) \quad (3)$$

Here, L denotes a Laplacian filter, and I_{bpn} and M are inputs; M is an $r \times r$ Laplacian mask, where r must be an odd number. The sum of all the elements in the mask should be zero. Laplacian filters extract edges according to differences in brightness. Because they react strongly to thin lines or points in an image, they are suitable for thresholding [24]:

$$I_{thn} = \begin{cases} 1, & \text{if } \tau_1 < I_{ln}(x, y) < \tau_2 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

Thresholding is the simplest method for segmenting images. Thresholding methods replace each pixel in an image with a black pixel if the pixel intensity is less than a fixed constant. To remove unnecessary areas after thresholding, areas with a small number of remaining pixels are removed. Laplacian filter detects only the edge located at the center of the changing area. Additionally, it is evident that thresholding the filtered image results in a narrower focus area. Therefore, we reconstructed the focal region as a morphological operation [25,26]:

$$I_{dn}(x, y) = I_{thn}(x, y) \oplus S_n \quad (5)$$

In a morphological operation, each pixel in an image is adjusted based on the values of other pixels in its neighborhood. Assume that the structural element for area dilation is defined as S_n ; if the structure element overlaps with a pixel in an input image, then the input image, I_{thn} is expanded. Figure 2 presents the results of the operations discussed above.

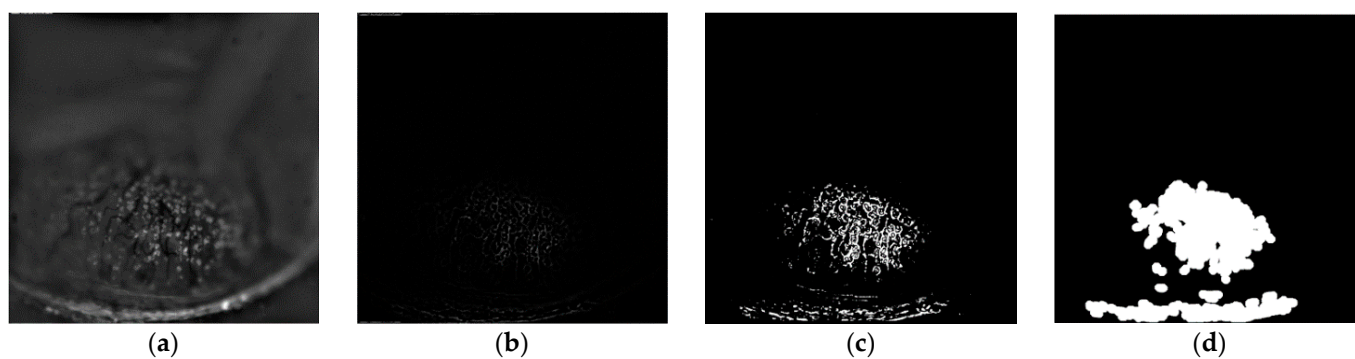


Figure 2. Results of local focus area detection. (a) Band-pass-filtered image, (b) Laplacian-filtered image, (c) thresholded image, and (d) dilated image.

2.5. Self-Adjusting Guided Filtered Image Fusion

The guided filter method was first proposed by He et al. [27]. A guided filter is a filter that preserves edges, e.g., a bilateral filter. A guided filter kernel is fast, regardless of its size and strength range, and is not impeded by a directional reversal structure. Guided filters were often used in image fusion in previous studies; thus, we optimized the filtering method to fit our algorithm:

$$I_{gni}(x, y) = a_k I_{dni} + b_k, \forall i \in r_k \quad (6)$$

$$r_k = 1 + (n - 1) \times s, n = 1, 2, 3 \dots m \quad (7)$$

$$m = \frac{1}{4} \times \text{image size} \quad (8)$$

A guided filter assumes that an output I_{gn} is a linear transformation of a guidance image I_{dn} in a window centered on a pixel k , where r_k is the window, and a_k and b_k are linear correlation coefficients that minimize the squared difference between an output image I_{gni} and input image I_{mi} :

$$I_{mi}(x, y) = I_n(x, y) \times I_{dn}(x, y) \quad (9)$$

$$E(a_k, b_k) = \sum_{i \in r_k} ((a_k I_{dni} + b_k - I_{mi})^2 + \epsilon a_k^2) \quad (10)$$

When the center pixel k changes, the result image I_{gn} also changes. In order to reduce this variation, the result image is determined by averaging the estimates from a_k and b_k .

The guided filter was utilized in a sliding window, and filters are applied to the target area according to the size of the window. However, it should be able to weight the image boundaries while preserving a wide area. To select an accurate focus area according to the microscope's field-of-view area, the window size was automatically adjusted with a self-adjusting guided filter [28]. Therefore, the guided, filtered image, which was affected by window size, was expanded to one-quarter of the size of the entire image, which accelerated the parameter adjustment process. The scale factor s determines the rate of expansion. Thus, we set factor $s = 2$ in the experiment. Figure 3 presents the results for various values of the window size r . If the r is set to a small value, a gap between the fused areas occurs. On the contrary, if the r is set too large, a fusion occurs with unnecessary parts of the image, making it impossible to create a natural all-in-focus image.

After multiplying the original image by the local focus area extraction mask, the focus areas obtained for each image were combined into a single all-in-focus image. Each image had a different focus area; therefore, different image sequence values were included. Additionally, for overlapping focus areas, we used the pixel-wise weighted averaging rule. The pixel-wise weighted averaging rule refers to the method of assigning weights to compensate for the brightness of images during the process of blending between pixels. The final focus area mask produced by the guided filter becomes blurred from the inside to the boundary lines, resulting in smaller pixel values. These pixel values are then regarded as weights. When the source images are fused with respect to the weights, smoothing results are obtained, while maintaining the boundaries between images. The procedure is shown in Algorithm 1.

Algorithm 1 Multi-focus image fusion algorithm.

```

1: Input  $I_N$ : Source images from fluorescence microscopies.
2: Output  $F$ , All-in-focus image.
3: //Obtain guided filtered focus map of source images
4: //Obtain output  $F$  by selecting the pixels  $(i, j)$  from the set of source images, which depends on
   the calculated weight of the guidance image  $I_{g_{ni}}$  for the respective pixels.
5: for  $i = 1 : p$ 
6:   for  $j = 1 : q$ 
7:      $I_{arm}(k) = \text{argsort}(I_{g_{ni}}(i, j))$ 
8:     //Arrange the calculated weights of the guidance image with respect to the source images.
9:     for  $k = 1 : N$  //where  $N$  is the number of source images to be fused
10:       $F(i, j) + = I_{arm}(k) \cdot I_k(i, j)$ 
11:    //Obtain output  $F$  by sequentially multiplying the source with the maximum weight.
12:    end for
13:  end for
14: end for

```

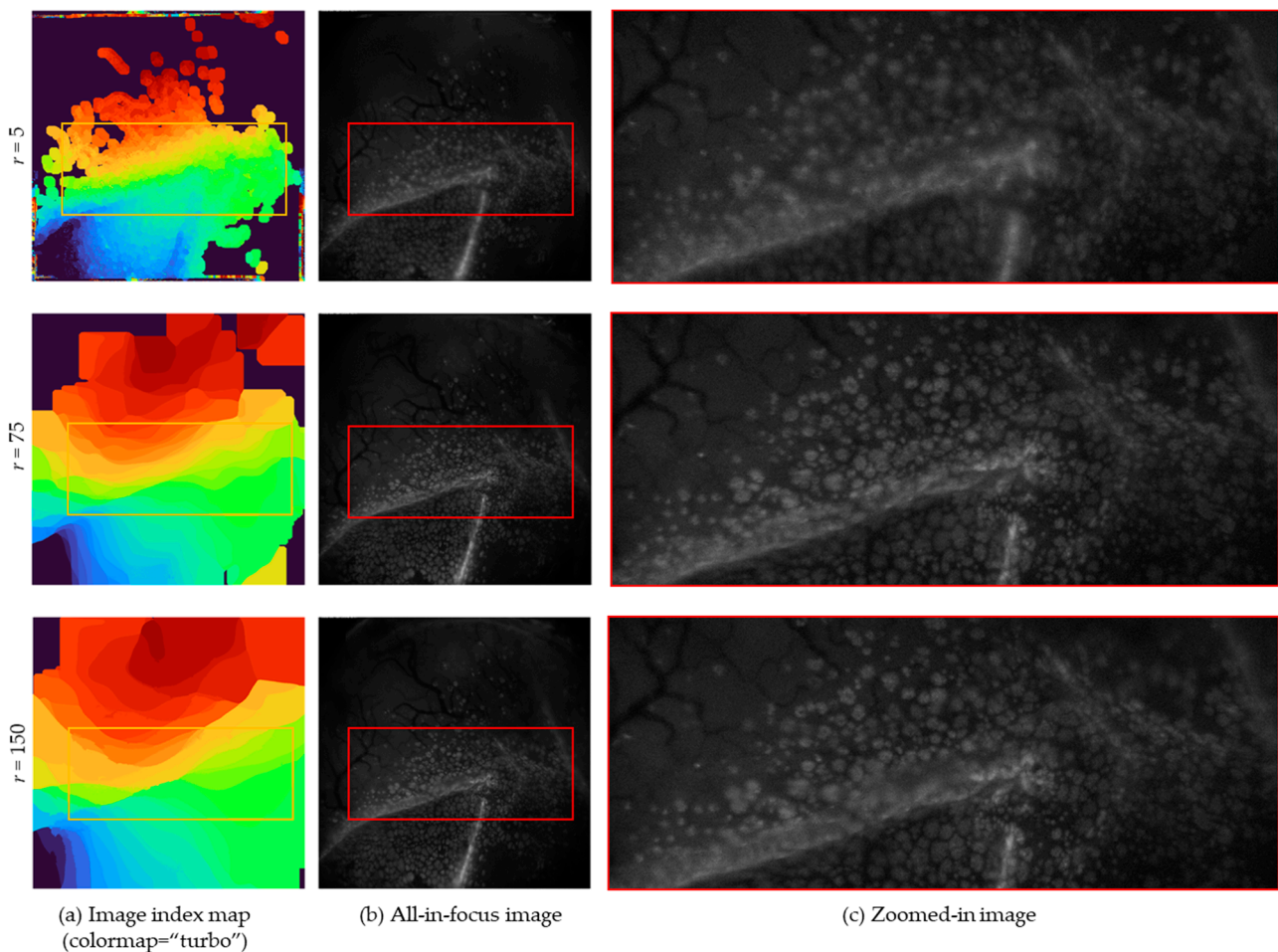


Figure 3. Image index map and processed results for different values of the window size r . Each column is separated by a r . (b) All-in-focus images that are fused based on image index maps in (a). The marked areas highlighted by the red box in (b) represent the zoomed-in images (c). By adjusting r , the area affected by the filter is also adjusted. If r is 5, as shown in the first column, it did not properly express the boundary features. In the third column, most areas in the image index map are indexed. Since information is extracted from a wide area, there is a disadvantage of obtaining information in an out-of-focus area. As shown in the second column, by choosing an appropriate r , a clear fusion result can be obtained without loss of features.

2.6. Objective Evaluation Metrics

An objective evaluation of fused images is difficult because there are no standard metrics for evaluating the image fusion process. “Full-reference” condition represents that reference image is secured, and there is a “no-reference” or “blind” condition where reference images are not available, as in many real applications. The image used in the first experiment is “Full-reference” condition, and the dataset used in the second experiment is “blind” condition [29]. Therefore, the following objective assessment metrics were applied according to the conditions.

First of all, there are the “Full-reference” state-only evaluation methods: Q_{MI} is an information-based convergence indicator based on a normalization that overcomes the instability of mutual-information-based indicators. It was proposed by Hossny et al. [30]:

$$Q_{MI} = 2 \left[\frac{MI(A, F)}{H(A) + H(F)} + \frac{MI(B, F)}{H(B) + H(F)} \right] \quad (11)$$

Here, $H(X)$ is the entropy of the image, and $MI(X, Y)$ is the mutual information value between two images, X and Y .

Q_{NCIE} is an information-based fusion indicator proposed by Wang et al. [31]; λ_i denotes the eigenvalues of a nonlinear correlation matrix:

$$Q_{NCIE} = 1 + \sum_{i=1}^3 \frac{\lambda_i}{3} \log_{256} \frac{\lambda_i}{3} \quad (12)$$

Q_G is the most-well-known image fusion evaluation metric that measures the degree of gradient information preserved in fused images relative to input images [32]:

$$Q_G = \frac{\sum_{i=1}^W \sum_{j=1}^H [Q^{AF}(i, j) \omega^A(i, j) + Q^{BF}(i, j) \omega^B(i, j)]}{\sum_{i=1}^W \sum_{j=1}^H (\omega^A(i, j) + \omega^B(i, j))} \quad (13)$$

Here, the width of the image is W , and the height is H ; $Q^{AF}(i, j) = Q_g^{AF}(i, j) Q_\lambda^{AF}(i, j)$, and Q_g^{AF} and Q_λ^{AF} are representative of the edge strength and gradient information preserved in the fused image relative to the original image, respectively. The same notation applies to Q^{BF} . ω^A and ω^B are the weights of Q^{AF} and Q^{BF} , respectively.

Q_p is an evaluation metric based on phase congruency. Phase congruency contains prominent feature information from images, such as edge and corner information [33]:

$$Q_p = (P_p)^\alpha (P_M)^\beta (P_m)^\gamma \quad (14)$$

Here, p , M , and m are the phase congruency, maximum moment, and minimum moment, respectively; P_p , P_M , and P_m are the maximum correlation coefficients between fused images and input images; and α , β , and γ are the parameters used to adjust the significance of each of the three coefficients, respectively.

Q_{CB} is a method based on the human visual system model. It consists of contrast filtering, local contrast calculation, contrast preservation, and quality guidance methods [34]:

$$Q_{GQM} = \lambda_A(i, j) Q_{AF}(i, j) + \lambda_B(i, j) Q_{BF}(i, j) \quad (15)$$

Here, Q^{AF} and Q^{BF} denote the contrast information of the input images preserved in a fused image, and λ denotes the weight value of an input image. Q_{CB} is defined as the mean value of Q_{GQM} as follows:

$$Q_{CB} = \overline{Q_{GQM}} \quad (16)$$

Peak signal-to-noise ratio ($PSNR$) and structural similarity index measure ($SSIM$) were used as full-reference quality assessment methods. $PSNR$ is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise

that affects the fidelity of its representation [35]. *PSNR* is most easily defined via the mean squared error (*MSE*). Given a noise-free $m \times n$ image and its noisy approximation, *PSNR* is defined as:

$$PSNR = 10 \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (17)$$

Here, MAX_I is the maximum possible pixel value of the image. Because it is measured in logarithmic scale, the unit is dB, and the smaller the loss, the higher the value. For lossless images, the *PSNR* is not defined because the *MSE* is zero.

The *SSIM* is used for measuring the similarity between two images [36]. *SSIM* is a perception-based model that considers image degradation as a perceived change in structural information, while incorporating important perceptual phenomena, including both luminance-masking and contrast-masking terms. The difference with other techniques such as *MSE* or *PSNR* is that these approaches estimate absolute errors. Given an original image and distorted image, *SSIM* is defined as:

$$SSIM(A, B) = \frac{(2\mu_A\mu_B + C_1)(2\sigma_{AB} + C_2)}{(\mu_A^2 + \mu_B^2 + C_1)(\sigma_A^2 + \sigma_B^2 + C_2)} \quad (18)$$

Here, μ_A is the average of A , σ_A^2 is the variance of A and the same notation applies to μ_B and σ_B^2 . σ_{AB} is the covariance of A and B . C_1 and C_2 are two variables to stabilize the division with the weak denominator.

No-reference methods were employed for fused images because reference images are commonly unavailable. One of the most representative no-referenced image quality assessments is BRISQUE, which was introduced by Mittal et al. [37]. BRISQUE is an algorithm that operates on the assumption that if a natural image is distorted, then the statistics of the corresponding image pixels is distorted. A natural image is an initial image captured by a camera that is not processed. Natural images exhibit regular statistical characteristics. The histogram of pixel values takes the form of a Gaussian distribution when processing the MSCN for such an image. For an image quality evaluation, after processing the MSCN, the pixels values were matched with a generalized Gaussian distribution (GGD) to utilize information regarding the pixel distribution as a characteristic feature. The parameters and variance values were compared to the GGD with the most similar forms to evaluate the characteristics of the target image.

Additionally, we defined the NIQE method. This method was also proposed by Mittal et al. [38]. The more similar the output of this method is to a test image, the better the quality of the test image. We also applied preprocessing using MSCN to divide images into patches. We could then derive BRISQUE characteristics within patches and calculate image quality values using mean vectors and covariance metrics.

3. Results and Discussion

In order to verify the proposed method with objective and subjective metrics, we compared its performance with some state-of-the-art methods, such as the discrete wavelet transform (DWT) image fusion, the quad-tree block-based image fusion [8], and Gaussian-filter-based multi-focus image fusion (GFDF) [37]. The first experiment assesses the “Full-reference” condition. We used Kaggle data science bowl 2018 datasets. Since this dataset aims to detect cell nuclei, the prepared data were acquired under various conditions and differed in imaging modalities. We selected two samples which were similar to our microscopic images and named them “Dark cell” and “Bright cell” (Figure 4). For the objective evaluation of the proposed method in this study, Gaussian blurring was applied to the original image to produce three blurred images. The second experiment evaluated the fusion performance by applying the appropriate metrics to the image set with the “blind” condition. The dataset used for the evaluation consisted of conjunctival goblet cell microscopic images taken from a mouse, consisting of 2048×2048 grayscale pixels. Each subset contained more than 20 images with different DOFs. All the experiments were implemented in MATLAB 2019a on an Intel i7-8700 CPU @ 3.20 GHz desktop with

32.00 GB RAM. The proposed method was compared to methods developed in previous works using program codes provided by the original authors [14,39].

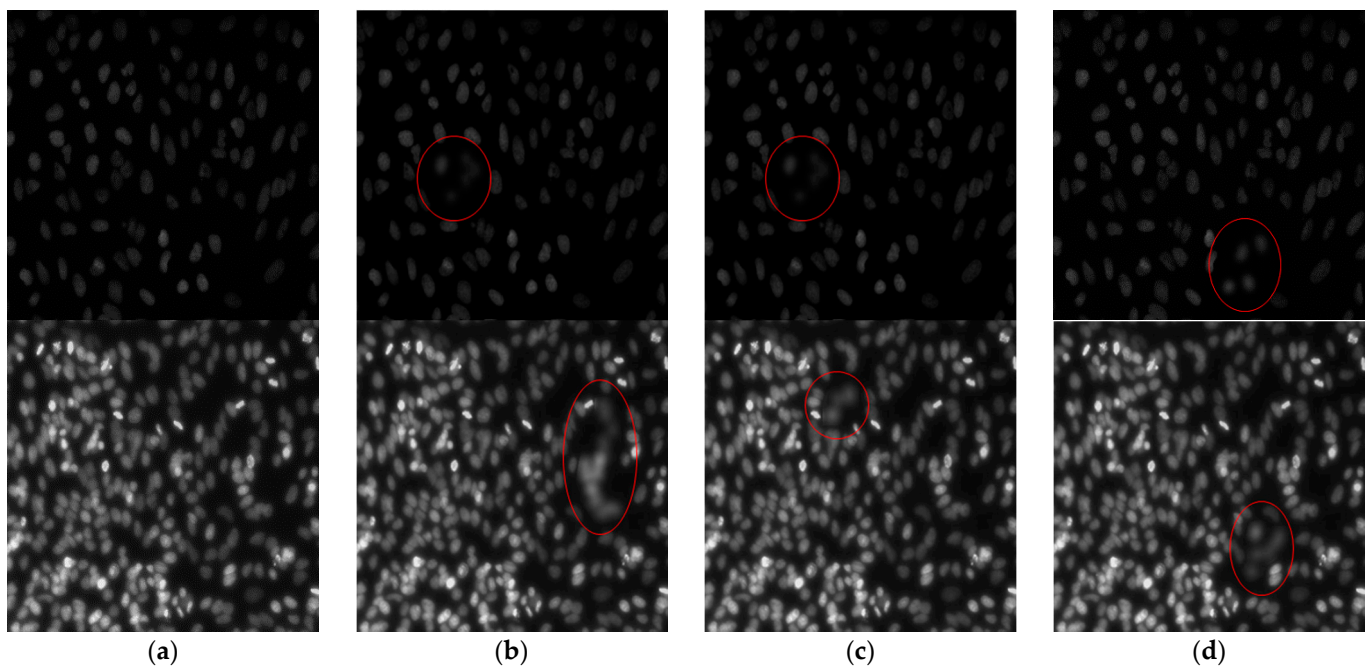


Figure 4. Grayscale image pairs for experiments with random blurring. The top row contains the “Dark cell” image set. The second row contains the “Bright cell” image set. (a) Original images and (b–d) randomly blurred images. From (b) to (d), we can see that some cells are defocused, and these are marked with a red circle.

Figure 5 presents the image fusion results for three images obtained using each MFIF method, and Figure 6 presents the details of the “Bright cell” fusion results. For the DWT and quad-tree methods, the boundaries of the focus areas remain in the resulting images; these areas are marked with a red rectangle in Figure 6. In the GFDF results, the boundaries of the focus areas are not visible, yet some details of the images are missing. The proposed method does not leave boundary lines in the focus areas, and the details of the images are preserved.

The image quality evaluations for the images presented in Figure 6 are listed in Tables 1 and 2. Comparing the objective metrics reveals that the images fused by the transform-domain methods lose gradient, structure, and edge information. To evaluate the images in the same way, Tables 1 and 2 are shown as the average of the evaluations of two images among the three input images. The GFDF method utilizes only the absolute difference between two images when detecting the focus area. Therefore, although it shows a high similarity in structure, other image fusion metrics are inferior to those of the other methods when there are more than three input images or overlapping focus areas. Regardless of the number of overlapping focal regions or local focal images, the proposed method is superior in terms of the amount of information loss, results of extracting edges, and quality of the fused results.

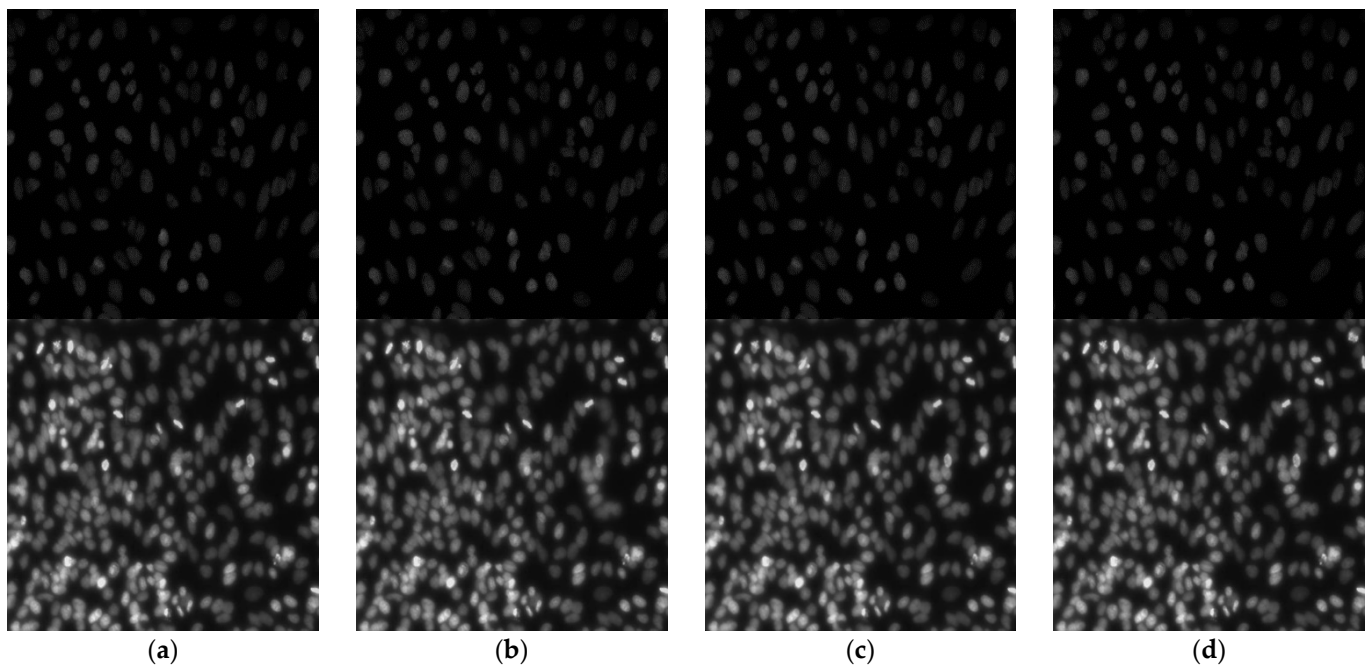


Figure 5. Fused image results generated by different methods. The top row contains the “Dark cell” image set. The bottom row contains the “Bright cell” image set. (a) DWT, (b) quad-tree, (c) GFDF, and the (d) proposed method.

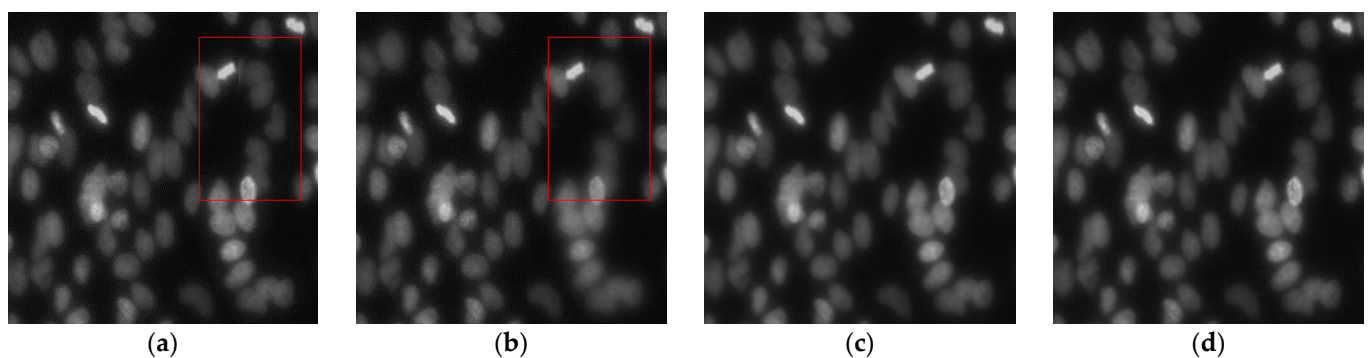


Figure 6. Details of “Bright cell” fused image results generated by different methods. (a) DWT, (b) quad-tree, (c) GFDF, and the (d) proposed method.

Table 1. Quantitative results for evaluation metrics on the “Dark cell” image with four other methods.

	QMI	QNCIE	QG	QP	QCB	PSNR	SSIM
DWT	1.8534	0.8322	0.9513	0.9363	0.9558	47.3966	0.9851
Quad-tree	1.7286	0.8298	0.9346	0.9064	0.9318	45.5006	0.9811
GFDF	1.4601	0.8242	0.8040	0.8489	0.8436	47.7980	0.9868
Ours	1.8796	0.8327	0.9519	0.9354	0.9726	47.4428	0.9854

Table 2. Quantitative results for evaluation metrics on the “Bright cell” image with four other methods.

	QMI	QNCIE	QG	QP	QCB	PSNR	SSIM
DWT	1.8902	0.9074	0.9630	0.9547	0.9629	43.7463	0.9860
Quad-tree	1.8083	0.9007	0.9593	0.9372	0.9467	42.8896	0.9856
GFDF	1.5250	0.8818	0.8919	0.8816	0.5276	43.3569	0.9874
Ours	1.9182	0.9098	0.9652	0.9578	0.9815	44.3206	0.9870

The second experiment was conducted to evaluate the performance of the proposed method on the “blind” condition image sets. Quad-tree and GFDF algorithms were implemented in two input images. DWT algorithm stated that it was able to fuse 13 images, but the provided code was composed for pair images. Thus, we conducted an experiment in the same conditions as Quad-tree and GFDF. On the contrary, the proposed method is mainly focused on merging more than 20 images at once. Figure 7 presents the image fusion results.

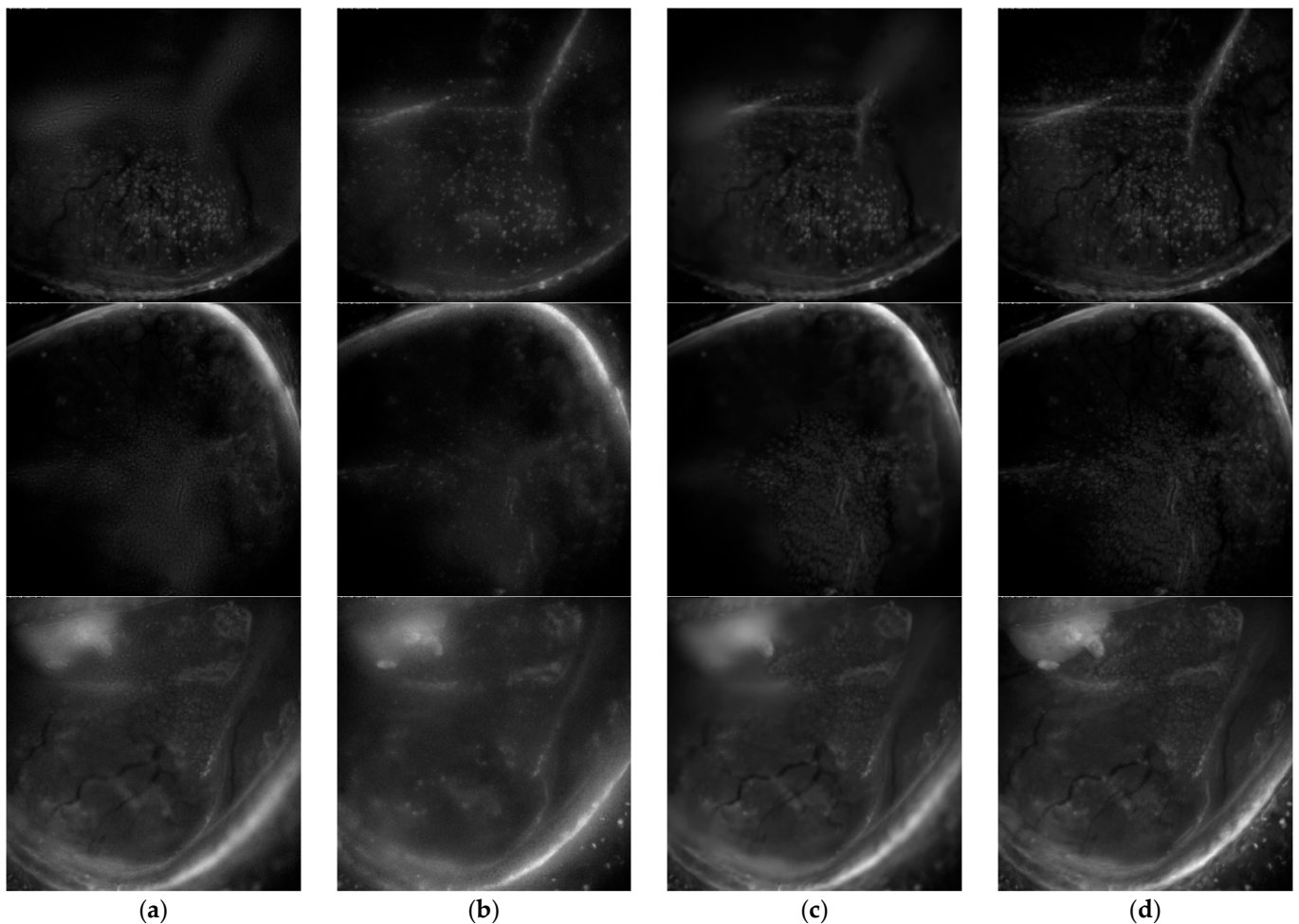


Figure 7. Conjunctival goblet cell, fused image results from different methods. (a) DWT, (b) quad-tree, (c) GFDF, and the (d) proposed method.

The blind image quality evaluation for the images presented in Figure 7 are listed in Tables 3–5. According to Tables 3–5, the no-reference image quality assessment indicates the results for each conjunctival goblet cell image. Blind reference-less image spatial quality evaluator (BRISQUE) methods sometimes performed better on GFDFs. However, the naturalness image quality evaluator (NIQE) measurements indicated that the proposed method yielded better results. In the case of the DWT method, one can observe that information loss during image reconstruction is unavoidable. This method suffers from a large amount of information loss in averaging blocks when it is applied to multiple source images. Unlike DWT, the quad-tree method automatically decomposes the window size according to the input image characteristics, but information is not preserved as the number of input images increases. In the case of GFDF, the differences between the images are used to detect focus areas. As the number of source images increases, only the source image information that is fused into the subsequent images is retained as the information from the initially generated focus areas disappears.

Table 3. Blind image quality evaluation results for first dataset of conjunctival goblet cell images.

	DWT	Quad-Tree	GFDF	Ours
BRISQUE	42.590	43.459	35.619	35.890
NIQE	12.969	12.515	3.442	2.955

Table 4. Blind image quality evaluation results for second dataset of conjunctival goblet cell images.

	DWT	Quad-Tree	GFDF	Ours
BRISQUE	42.239	43.443	36.941	31.006
NIQE	11.768	12.063	3.883	2.855

Table 5. Blind image quality evaluation results for third dataset of conjunctival goblet cell images.

	DWT	Quad-Tree	GFDF	Ours
BRISQUE	42.668	43.499	19.908	31.676
NIQE	15.775	19.467	4.221	3.28

From a quantitative perspective, Table 4 indicates that the performance of the proposed method is superior, regardless of the number of source images. The results obtained by the proposed method are more stable and systematic than those of the other fusion methods in terms of the objective evaluation metrics.

4. Conclusions

In this work, we presented a multi-focus image fusion method applied to a large quantity of conjunctival microscopic images. Wide-field fluorescence microscopy acquired multiple images with the axial translation of focus, and the large quantity of images transformed into single all-in-focus images through multi-focus image fusion. The proposed method is highly effective in that it performs fusion without being affected by the size and noise of the input image and the number of source images. The proposed method uses the high-frequency characteristics of the focal area to determine the area with a Laplacian filter. Nevertheless, the focus region is detected using the Laplacian filter, and there may be some undetectable parts due to ambiguous boundaries. In addition, the Laplacian filter captures the center of the focus region; we used a morphological operation to compensate for this. The proposed method works on the basis of fixed structural elements, where it is difficult to completely reconstruct the desired area.

However, the experiment was carried out in order to image a live animal model, and the proposed method showed several advantages over previous MFIF methods. First, it prevented visible artifacts such as block shapes and blurring. Additionally, regardless of the number of source images, it was confirmed that an image could be fused using just one iteration and that the proposed method was robust to images with noise. Because differences between microscopic images and general images appear when defining thresholds following a Laplacian transformation, it would be useful to investigate how to select the appropriate thresholds according to the target images. Additionally, developing a better method for the focus area detection is worth additional consideration.

Image fusion techniques are commonly applied in various fields, such as digital photography and medical diagnosis. In particular, it is important that optical microscopic image fusion be performed without losing information. It is expected that both experts and non-experts will be able to fuse images easily using the proposed algorithm.

Author Contributions: Conceptualization, K.H.K. and S.Y.; methodology, J.L. (Jiyoung Lee) and S.J.; software, J.L. (Jiyoung Lee) and S.J.; validation, J.L. (Jiyoung Lee) and S.J.; formal analysis, J.L. (Jiyoung Lee); investigation, J.L. (Jiyoung Lee), S.J. and J.S.; resources, J.L. (Jungbin Lee), T.K. and S.K.; data curation, J.L. (Jiyoung Lee) and J.L. (Jungbin Lee); writing—original draft preparation, J.L. (Jiyoung Lee) and S.J.; writing—review and editing, S.Y., K.H.K. and J.S.; visualization,

J.L. (Jiyoung Lee) and S.J.; supervision, S.Y. and K.H.K.; project administration, S.Y. and K.H.K.; funding acquisition, S.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported partly by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education (MSIT) (No. 2019R1F1A1058971), and partly by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (No. 2020-0-00989).

Institutional Review Board Statement: All the animal experimental procedures were approved by the Institutional Animal Care & Use Committee at the Pohang University of Science and Technology (IACUC, approval number: POSTECH-2015-0030-R2) and were conducted in accordance with the guidelines.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Li, S.T.; Kang, X.D.; Fang, L.Y.; Hu, J.W.; Yin, H.T. Pixel-level image fusion: A survey of the state of the art. *Inf. Fusion* **2017**, *33*, 100–112.
- Ralph, R.A. Conjunctival goblet cell density in normal subjects and in dry eye syndromes. *Investig. Ophthalmol.* **1975**, *14*, 299–302.
- Colorado, L.H.; Alzahrani, Y.; Pritchard, N.; Efron, N. Assessment of conjunctival goblet cell density using laser scanning confocal microscopy versus impression cytology. *Contact Lens Anterior Eye* **2016**, *39*, 221–226. [[PubMed](#)]
- Cinotti, E.; Singer, A.; Labeille, B.; Grivet, D.; Rubegni, P.; Douchet, C.; Cambazard, F.; Thuret, G.; Gain, P.; Perrot, J.L. Handheld in vivo reflectance confocal microscopy for the diagnosis of eyelid margin and conjunctival tumors. *JAMA Ophthalmol.* **2017**, *135*, 845–851. [[CrossRef](#)] [[PubMed](#)]
- Lee, J.; Kim, S.; Yoon, C.H.; Kim, M.J.; Kim, K.H. Moxifloxacin based axially swept wide-field fluorescence microscopy for high-speed imaging of conjunctival goblet cells. *Biomed. Opt. Express* **2020**, *11*, 4890–4900. [[CrossRef](#)]
- Bhat, S.; Koundal, D. Multi-focus image fusion techniques: A survey. *Artif. Intell. Rev.* **2021**, *54*, 5735–5787. [[CrossRef](#)]
- Kaur, H.; Koundal, D.; Kadyan, V. Image fusion techniques: A survey. *Arch. Comput. Methods Eng.* **2021**, 1–23. [[CrossRef](#)]
- Li, H.; Manjunath, B.S.; Mitra, S.K. Multisensor Image Fusion Using the Wavelet Transform. *Graph. Models Image Process.* **1995**, *57*, 235–245. [[CrossRef](#)]
- Rockinger, O. Image sequence fusion using a shift-invariant wavelet transform. In Proceedings of the International Conference on Image Processing, Santa Barbara, CA, USA, 26–29 October 1997; IEEE: Manhattan, NY, USA, 1997; pp. 288–291.
- Mitianoudis, N.; Stathaki, T. Pixel-based and region-based image fusion schemes using ICA bases. *Inf. Fusion* **2007**, *8*, 131–142. [[CrossRef](#)]
- Tang, J.S. A contrast based image fusion technique in the DCT domain. *Digit. Signal Process.* **2004**, *14*, 218–226. [[CrossRef](#)]
- Zhang, Q.; Liu, Y.; Blum, R.S.; Han, J.G.; Tao, D.C. Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: A review. *Inf. Fusion* **2018**, *40*, 57–75. [[CrossRef](#)]
- Li, S.; Kwok, J.T.; Wang, Y. Combination of images with diverse focuses using the spatial frequency. *Inf. Fusion* **2001**, *2*, 169–176. [[CrossRef](#)]
- Bai, X.Z.; Zhang, Y.; Zhou, F.G.; Xue, B.D. Quadtree-based multi-focus image fusion using a weighted focus-measure. *Inf. Fusion* **2015**, *22*, 105–118. [[CrossRef](#)]
- 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](#)]
- Huang, Y.; Li, W.S.; Gao, M.L.; Liu, Z. Algebraic Multi-Grid Based Multi-Focus Image Fusion Using Watershed Algorithm. *IEEE Access* **2018**, *6*, 47082–47091. [[CrossRef](#)]
- Bhat, S.; Koundal, D. Multi-focus Image Fusion using Neutrosophic based Wavelet Transform. *Appl. Soft Comput.* **2021**, *106*, 107307. [[CrossRef](#)]
- Yang, Y.; Zhang, Y.M.; Wu, J.H.; Li, L.Y.; Huang, S.Y. Multi-Focus Image Fusion Based on a Non-Fixed-Base Dictionary and Multi-Measure Optimization. *IEEE Access* **2019**, *7*, 46376–46388. [[CrossRef](#)]
- Xu, K.P.; Qin, Z.; Wang, G.L.; Zhang, H.D.; Huang, K.; Ye, S.X. Multi-focus Image Fusion using Fully Convolutional Two-stream Network for Visual Sensors. *KSII Trans. Internet Inf. Syst.* **2018**, *12*, 2253–2272.
- Bracewell, R.N.; Bracewell, R.N. *The Fourier Transform and Its Applications*; McGraw-Hill: New York, NY, USA, 1986; Volume 31999.
- 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](#)]
- Nayar, S.K.; Nakagawa, Y. Shape from focus. *IEEE Trans. Pattern Anal. Mach. Intell.* **1994**, *16*, 824–831. [[CrossRef](#)]
- Burt, P.J.; Adelson, E.H. The Laplacian Pyramid as a Compact Image Code. *IEEE Trans. Commun.* **1983**, *31*, 532–540.
- Sezgin, M.; Sankur, B. Survey over image thresholding techniques and quantitative performance evaluation. *J. Electron. Imaging* **2004**, *13*, 146–168.

25. Haralick, R.M.; Sternberg, S.R.; Zhuang, X.H. Image-Analysis Using Mathematical Morphology. *IEEE Trans. Pattern Anal. Mach. Intell.* **1987**, *9*, 532–550. [[CrossRef](#)] [[PubMed](#)]
26. De, I.; Chanda, B. Multi-focus image fusion using a morphology-based focus measure in a quad-tree structure. *Inf. Fusion* **2013**, *14*, 136–146.
27. He, K.; Sun, J.; Tang, X. Guided image filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 1397–1409.
28. Li, H.; Qiu, H.; Yu, Z.; Li, B. Multifocus image fusion via fixed window technique of multiscale images and non-local means filtering. *Signal Process.* **2017**, *138*, 71–85. [[CrossRef](#)]
29. Liu, Z.; Blasch, E.; Xue, Z.Y.; Zhao, J.Y.; 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](#)]
30. Hossny, M.; Nahavandi, S.; Creighton, D. Comments on ‘Information measure for performance of image fusion’. *Electron. Lett.* **2008**, *44*, 1066–1067. [[CrossRef](#)]
31. Wang, Q.; Shen, Y.; Zhang, J.Q. A nonlinear correlation measure for multivariable data set. *Phys. D Nonlinear Phenom.* **2005**, *200*, 287–295. [[CrossRef](#)]
32. Xydeas, C.A.; Petrovic, V. Objective image fusion performance measure. *Electron. Lett.* **2000**, *36*, 308–309.
33. Zhao, J.Y.; 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.
34. Chen, Y.; Blum, R.S. A new automated quality assessment algorithm for image fusion. *Image Vis. Comput.* **2009**, *27*, 1421–1432. [[CrossRef](#)]
35. Huynh-Thu, Q.; Ghanbari, M. Scope of validity of PSNR in image/video quality assessment. *Electron. Lett.* **2008**, *44*, 800–801. [[CrossRef](#)]
36. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)]
37. Mittal, A.; Moorthy, A.K.; Bovik, A.C. No-Reference Image Quality Assessment in the Spatial Domain. *IEEE Trans. Image Process.* **2012**, *21*, 4695–4708. [[CrossRef](#)] [[PubMed](#)]
38. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a “completely blind” image quality analyzer. *IEEE Signal Process. Lett.* **2012**, *20*, 209–212.
39. Qiu, X.H.; Li, M.; Zhang, L.Q.; Yuan, X.J. Guided filter-based multi-focus image fusion through focus region detection. *Signal Process.-Image Commun.* **2019**, *72*, 35–46. [[CrossRef](#)]