

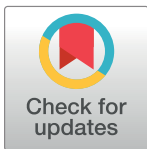
RESEARCH ARTICLE

A multi-focus image fusion method via region mosaicking on Laplacian pyramids

Liang Kou¹, Liguang Zhang^{1*}, Kejia Zhang¹, Jianguo Sun¹, Qilong Han¹, Zilong Jin²

1 College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China, **2** School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing 210044, China

* zhangliguo@hrbeu.edu.cn



Abstract

In this paper, a method named Region Mosaicking on Laplacian Pyramids (RMLP) is proposed to fuse multi-focus images that is captured by microscope. First, the Sum-Modified-Laplacian is applied to measure the focus of multi-focus images. Then the density-based region growing algorithm is utilized to segment the focused region mask of each image. Finally, the mask is decomposed into a mask pyramid to supervise region mosaicking on a Laplacian pyramid. The region level pyramid keeps more original information than the pixel level. The experiment results show that RMLP has best performance in quantitative comparison with other methods. In addition, RMLP is insensitive to noise and can reduce the color distortion of the fused images on two datasets.

OPEN ACCESS

Citation: Kou L, Zhang L, Zhang K, Sun J, Han Q, Jin Z (2018) A multi-focus image fusion method via region mosaicking on Laplacian pyramids. PLoS ONE 13(5): e0191085. <https://doi.org/10.1371/journal.pone.0191085>

Editor: A Lenin Fred, Mar Ephraem College of Engineering & Technology, INDIA

Received: April 2, 2017

Accepted: December 12, 2017

Published: May 17, 2018

Copyright: © 2018 Kou et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work is supported by the National Science Foundation of China (Grant No. 61501132, Grant No. 61370084), the China Postdoctoral Science Foundation No. 2016M591515, the Heilongjiang Postdoctoral Sustentation Fund with No. LBH-Z14055, and Harbin Application Technology Research and Development Project (Grant No. 2016RAQXJ063, Grant No. 2016RAXXJ013).

Introduction

Because of limited depth of field, high magnification optical cameras, such as microscopes or macro-photography, cannot capture an object that is totally in focus. When capturing the object/scene with the camera, typically only one or a few small regions of the image lying in focus are clear. Multi-focus images fusion is a common remedy to solve this problem. A set of images with different focuses are captured, then fused to produce an 'all-in-focus' image that is clear everywhere [1] [2] [3] [4] [5] [6]. The process of synthesizing the all-in-focus image is called multi-focus image fusion. Multi-focus images fusion has been proven valuable in many applications such as microscope imaging [7], image deblurring [8], shape from focus [7] [9] and information forensics [10].

According to fusion methods, the multi-focus images fusion can be categorized into two types: transform domain based and spatial domain based methods. In the former one, all the source images are firstly transformed from spatial domain to transform domain by using Fourier transform, multi-scale decomposition or other methods. After being combined by certain guidelines, the fused coefficients are inversely transformed back to spatial domain to obtain the expected fusion image. Many transform domain based algorithms are brought forward, such as the Laplacian pyramid (LP) [11], the Discrete Wavelet Transform (DWT) [12] [13], the gradient pyramid [14] [15], the contrast pyramid (CP) [16], the ratio-of-low pass pyramid

Competing interests: The authors have declared that no competing interests exist.

[17], the Shift-invariant DWT (SIDWT) [18], the Complex Wavelet Transform (CWT) [19] [20] and the Contourlet Transform [21] [22] [23].

Unlike transform domain based ones, the spatial domain based methods refer to the strategies that the source images are fused directly in gray space. These algorithms include weighted average fusion, weighted linear fusion, Principal Component Analysis (PCA) [24], Neural Network, etc. The spatial domain based multi-focus image fusion can also be divided into two categories: pixel based and block based methods. In the pixel based methods, pixels of source images are fused directly after exactly matching source images. The pixel based fusion methods can significantly improve the visual effects of the fused image, however, it is sensitive to noise and misregistration [25] [26]. When applying the block based methods, source images are firstly divided into $N \times M$ blocks. After comparing the sharpness of the blocks at corresponding position, the blocks with higher sharpness are spliced together to produce an all-in-focus image [16]. This kind of methods may result in block artifacts. In addition, the size of blocks can significantly affect the image fusion quality. Compared with the spatial domain based ones, the transform domain based methods can produce better fusion effect, whose limitations are the high cost and complexity of its computation.

As a representative of transform domain methods, the pyramid based methods have already been researched widely. Since these methods belong to the pixel level image fusion in a broad sense, most of them are sensitive to noise. Noise pixels and sharp regions of image are both high frequency signal, so it is hard to distinguish them by their frequency bands. The noises can be falsely identified as the pixels on the focal plane. The transform domain based methods apply global information to produce all-in-focus image, therefore a small change (is caused by noise) of any coefficient in the transformed domain may lead to changes of all pixels in spatial domain [27]. To solve the problem of noisy sensitivity, gradient map filtering (a directional change in the intensity or color in an image) [14] and multiple coefficient selection principles [28] are proposed, however, their performance depends on fine-tuned parameters.

For image fusion, the weighted approaches are more intuitive than pyramid-based methods [29] [30] [31]. When using weighted approaches, each pixel is assigned a weight which is calculated based on the focus degree of the source image. Fused image is the weighted summation of all corresponding pixels. Region mosaicking is a special weighted linear fusion, which sets one to the weight of the pixels lying at the focusing region, zero to the weight of the other pixels. In [32] Agarwala *et al.* proposed an 'iterative digital photomontage' method, which uses an interactive framework to combine source images into a single image. In the process, the graph cut algorithm is employed to segment the optimized mask. Mosaic algorithms often introduce block artifacts, while it can preserve original information of source images. In recent years, the sparse reconstruction methods are utilized to fuse weighted multi-focus images [10] [27] [29]. Sparse coefficients with the over complete dictionary are used to represent the multi-focus images. The coefficients are then combined with a choose-max rule, from which the fused image is reconstructed with respect to the over-complete dictionary. In the experiment results, the sparse weighted fusion achieved the highest performance among the existing weighted linear approaches. Recently, an artificial neural network model, pulse coupled neural network (PCNN), developed by Eckhorn *et al.* [33] has been employed in many applications of image fusion including weighted linear fusion [1] [34] [35] [36]. PCNN can automatically measure focus on the source images and then adjust weights of pixels. This method exhibits good performances in both visual effect and objective evaluation criteria, but a large number of its parameters often results in computational inefficiency.

When observing micro object by microscope, the characteristic of the captured multi-focus images is that the focus regions of adjacent frames are continuous, as shown in Fig 1. The

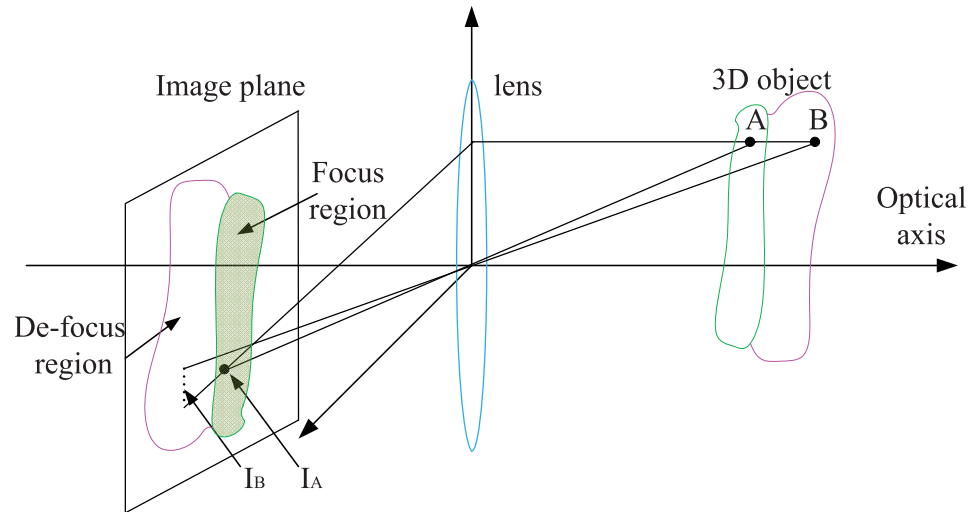


Fig 1. An illustration of 3D object imaging with an optical camera. I_A and I_B are image pixels of point A and B respectively. With the current focal length, the surface that point A lies on is focused while that of point B is not. It can be seen that in-focus pixels in the image plane form a continuous region. By adjusting the object distance to the lens, a series of defocused (part-in-focus) images could be obtained.

<https://doi.org/10.1371/journal.pone.0191085.g001>

traditional fusion methods devote attention to discrete pixels on focus plane and do not fully utilize the continuity. For instance, when using pyramid method (such as Laplacian pyramid, contrast pyramid and gradient pyramid) to fuse a set of multi-focus images, a pixel and its neighbors often belong to different layers of pyramid. In this case, corresponding focally regions of different layers are not similar multiscale shapes, which suggest that a great many pixels of clear regions in source images are lost. In the final fused image fusion, errors and distortion may be produced because of the above problems. In [37], Hariharan *et al.* use adaptive segmentation of focally connected regions to synthesize an all-in-focus image. Their method employs overlapping focal regions to extend the depth of field, while retaining the visual verisimilitude of the scene.

This paper proposes a simple and effective approach, Region Mosaicking on Laplacian Pyramids (hereafter referred to as RMLP for short), to fuse multi-focus images. It is based on the observation that the in-focus pixels in a multi-focus image form continuous regions. RMLP uses Density-Based Region Growing (DBRG) to generate a focus region mask for all of the multi-focus images. In DBRG, both regions growing and regions filtering are used to identify appropriate focus regions and reduce the impact of noises. A segmented focused region mask is decomposed into a mask pyramid, which is then applied to supervise the region mosaicking on a Laplacian pyramid. RMLP also improves pixel level pyramid fusion at the region level, where the imaging characteristics of multi-focus images are utilized and the continuity of segmented focused regions is incorporated. In RMLP, decomposition values of a pixel at corresponding position of different pyramid layers are taken from the exactly same multi-focus image, which guarantees that distortion artifacts are reduced to minimum. In addition, RMLP can also significantly reduce the artifacts introduced by weighted linear fusion approaches.

The remainder of this paper is organized as follows. An overview of the proposed approach is introduced in Section 2. Then the focus region segmentation is described in Section 3. The image fusion procedure of RMLP is given in Section 4. After providing experimental results, the paper is concluded in Section 6.

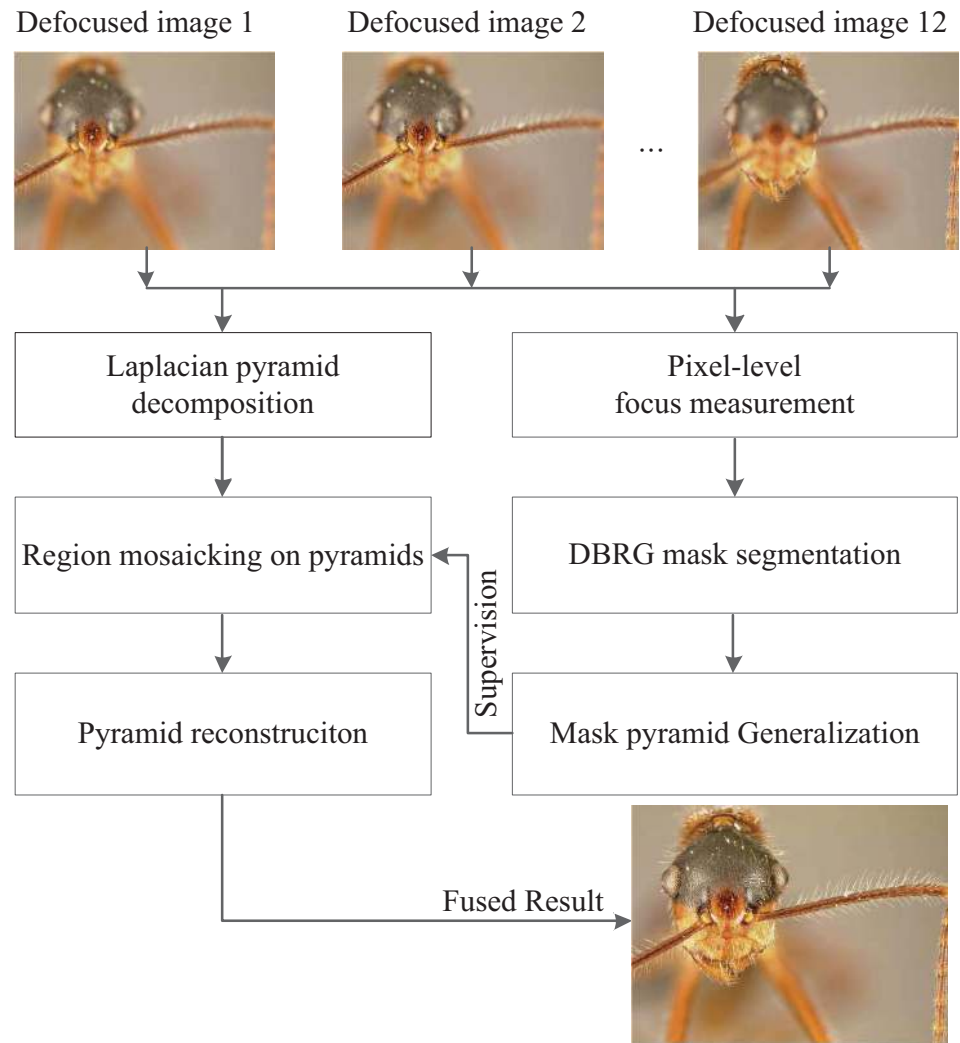


Fig 2. Flowchart of multi-focus images fusion based on RMLP approach.

<https://doi.org/10.1371/journal.pone.0191085.g002>

Overview of RMLP approach

The flowchart of RMLP is shown in Fig 2: the competition on Sum-Modified-Laplacian (SML) is used firstly to detect in-focus (clear) pixels in each image. The SML measurement is employed as the dominant cue for fusion, which guarantees that the information from multi-focus images is preserved. Fig 2 shows only three defocused images (the total number is 12) as the limitation of space.

DBRG is then applied to refine the extracted clear pixels, connect them into clear regions and form a focus region mask (Section 3). The focus region mask is then decomposed into a mask pyramid, corresponding to the Laplacian pyramids of multi-focus images. The mask pyramid contains label values of multi-layers, indicating which pyramid will be selected in the fusion procedure (Section 4).

Since the mask image contains continuous regions instead of disconnected pixels, in the fusion procedure, one segmented region will be selected from each Laplacian pyramid image so that all the selected regions cover the pyramid. This procedure is referred as region

mosaicking on pyramids. After region mosaicking, a reconstruction procedure is carried out to obtain the final ‘all-in-focus’ image as the output.

Focus region segmentation

Focus measurement based on Sum-Modified-Laplacian

For multi-focus image fusion, focus measurement needs to be done before focus region segmentation and image fusion. Therefore choosing an appropriate measurement method is crucial for subsequent process. In [38], Huang *et al.* contrastively analyzed several typical measurement methods, including Energy of image Gradient (EOG), Energy of Laplacian (EOL) of the image, Sum-Modified-Laplacian (SML), and spatial frequency (SF). In their experimental results, SML has the best performance in the aspect of image quality. Performance of EOL slightly worse than SML, but it is more computational efficient. For higher quality of fused image, SML is used to measure focus by RMLP. SML is presented by Nayar *et al.* in 1994 [39]. They firstly defined a modified Laplacian as

$$\nabla_M^2 I(x, y) = \left| \frac{\partial^2 I}{\partial x^2} \right| + \left| \frac{\partial^2 I}{\partial y^2} \right|. \tag{1}$$

Where $I(x, y)$ is a source image. In order to accommodate for possible variations in the size of texture elements, a variable spacing (*step*) between the pixels is used to compute the derivatives. In this paper, the difference between adjacent pixels is used to replace differential coefficient, i.e., *step* equals to one. The discrete approximation of Eq 1 is calculated as

$$\begin{aligned} \nabla_M^2 I(x, y) = & |2I(x, y) - I(x - 1, y) - i(x + 1, y)| \\ & + |2I(x, y) - I(x, y - 1) - i(x, y + 1)| \end{aligned} \tag{2}$$

A small window of size $(-w, w)$ is defined around a pixel (i, j) , then compute the focus at the pixel (i, j) as the sum of modified Laplacian:

$$SML = \sum_{x=i-w}^{i+w} \sum_{y=j-w}^{j+w} \nabla_M^2 I(x, y), \text{ for } \nabla_M^2 I(x, y) \geq T. \tag{3}$$

Where T is a discrimination threshold. In this paper, the size of windows is empirically set 3×3 , i.e., $w = 1$.

Focus region segmentation based on region growing

Focus regions have high sharpness, whose SML is much larger than the one in defocus regions. A mask image M_0 is defined whose pixel values vary in $[1, N]$. The pixel-level mask image M_0 is initially labeled with

$$M_0(i, j) = \arg \max_n |SML_n(i, j)|, n \in [1, N]. \tag{4}$$

Where $SML_n(i, j)$ is the SML of the n^{th} multi-focus image at pixel (i, j) .

Image segmentation is applied in many scenarios as a common pretreatment method [40] [41]. The idea of Density-Based Region Growing (DBRG) image segmentation is that the pixels with same label are represented as a cluster. The density distribution of clusters is then analyzed. The spatial neighborhood $\Omega(x, y)$ of a given pixel (i, j) is defined as a circle centered at the pixel with radius R , where R is determined experimentally as shown in Fig 3 The density

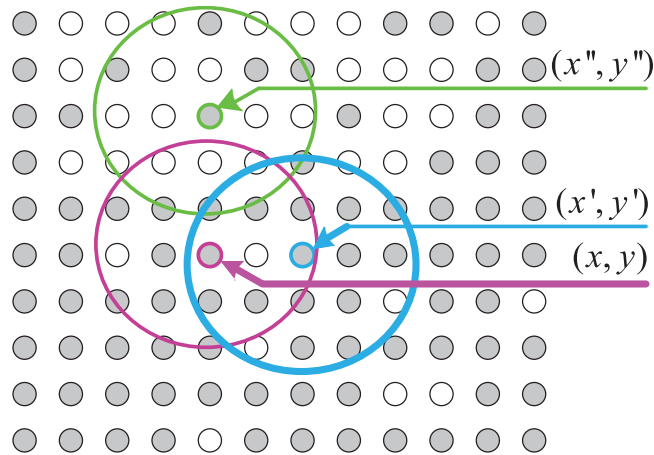


Fig 3. An illustration of density-connectivity with same mask label.

<https://doi.org/10.1371/journal.pone.0191085.g003>

distribution in $\Omega(x, y)$ is defined as

$$D_{\Omega(x,y)}(n) = \left(\frac{1}{\pi R^2} \sum_{(i,j) \in \Omega(x,y)} \delta(M_0(i,j) = n) \right). \tag{5}$$

Where the Boolean function $\delta(\cdot)$ is defined as

$$\delta(x) = \begin{cases} 1, & x \text{ is true} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

If the maximum density $\max_n (D_{\Omega(x,y)}(n))$ of $\Omega(x, y)$ is larger than a threshold 0.5, then the pixel (i, j) is called a seed pixel and $\Omega(x, y)$ forms a seed region, as shown by the gray pixels (circles) in Fig 3. In Fig 3 the pixel (x', y') is density-connected with pixel (x, y) if (x', y') is within the spatial neighborhood (mask) of (x, y) . Since the pixel (x'', y'') does not belong to the spatial neighborhood, it is not the density-connected with (x, y) according to the density-connectivity defined above.

As a preprocessing of multi-focus image fusion algorithm, DBRG is to find clear focus regions with high sharpness, so segmented region must be close to the seed point and the smaller the segmented region the better. Therefore we employ simple “checker-board” distance as the distance metric. The “checker-board” distance is the most widely used metric in image segmentation. Euclidean distance is often used to express similarity measure between feature vectors of image. The topological distance can be used to segment maximum connected region of image [42]. However, considering too much noise will be introduced into the focus regions. DBRG segmentation algorithm takes a set of mask pixels as input and output a set of mask regions. The algorithm is presented in Algorithm 1.

Algorithm 1: Mask Segmentation

Input: Pixel-level mask image M_0

Output: Region-level mask image R

- 1 Set all of the pixels of M_0 to unlabeled;
- 2 Search the unlabeled pixels in M_0 to find seed pixels and seed regions;
- 3 **for** each seed pixel (x, y) and seed region $\Omega(x, y)$ **do**
- 4 Create a new cluster with cluster label $R(x,y) = \arg \max_n (D_{\Omega(x,y)}(n))$

```

5   for each unlabeled pixel (x', y') density-connected with (x, y) do
6       Add (x', y') to the same cluster and mark (x', y') with label
           R(x, y);
7   end
8   end
9   for each unlabeled pixel (x'', y'') do
10      //noise pixels or pixels from smooth region.
11      Choose the labeled pixel most adjacent to (x'', y''); //let the
           pixel be (x, y).
12      Add (x'', y'') to the cluster of (x, y) and mark it with the
           corresponding label;
13   end
14 Create a region-level mask image R based on clusters;

```

The density-connectivity among pixels is transitive due to density reachability, which is consistent with the imaging characteristics of multi-focus images. In defocus regions, focus measurements are not stable since the SML is small. Noisy pixels can have larger SML than the true focused pixels. The density based thresholding and growing can reduce the label errors on smooth regions or noisy pixels.

Image fusion

After segmenting the focus region, the mask is generated with focused regions of source images. The mask is used to build the mask pyramid and supervise the fusion process of multi-focus images. This procedure is called region mosaicking on pyramid.

Pyramid based fusion

The initial layer of Gaussian pyramid, G_0 , is the source image $I(x, y)$. In [16], the k^{th} layer of Gaussian pyramid can be expressed as:

$$G_k(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) G_{k-1}(2i + m, 2j + n), \tag{7}$$

$$k \in [1, K], i \in [0, R_k], j \in [0, C_k]$$

Where $w(m, n)$ is a window function with Gaussian low-pass filter, K is the number of layers, R_k and C_k are the number of columns and rows in the k^{th} layer pyramid respectively. The bottom pyramid G_0 constitute the Gaussian pyramid with other pyramids G_1, \dots, G_{K-1} .

Laplacian pyramid can be obtained through calculating difference between two adjacent layers of Gaussian pyramid. It is defined as

$$\begin{cases} LP_k = G_k - G_{k+1}^*, k \in [0, K) \\ LP_k = G_k, k = K \end{cases} \tag{8}$$

G_{k+1}^* is expanded G_{k+1} , whose size is same with G_k .

The classical Laplacian pyramid (LP) algorithm exploits a pixel level competition and fusion on pyramids as

$$\begin{cases} F_k(x, y) = LP_k^n(x, y), k \in [0, K) \\ \hat{n} = \arg \max_n |LP_k^n(x, y)|, n \in [1, N] \end{cases} \tag{9}$$

Where F_k denotes the k^{th} layer of the fused LP, $LP_k^n(x, y)$ denotes LP value of pixel (x, y) at k^{th} layer of the n^{th} multi-focus image and N denotes the total number of multi-focus images.

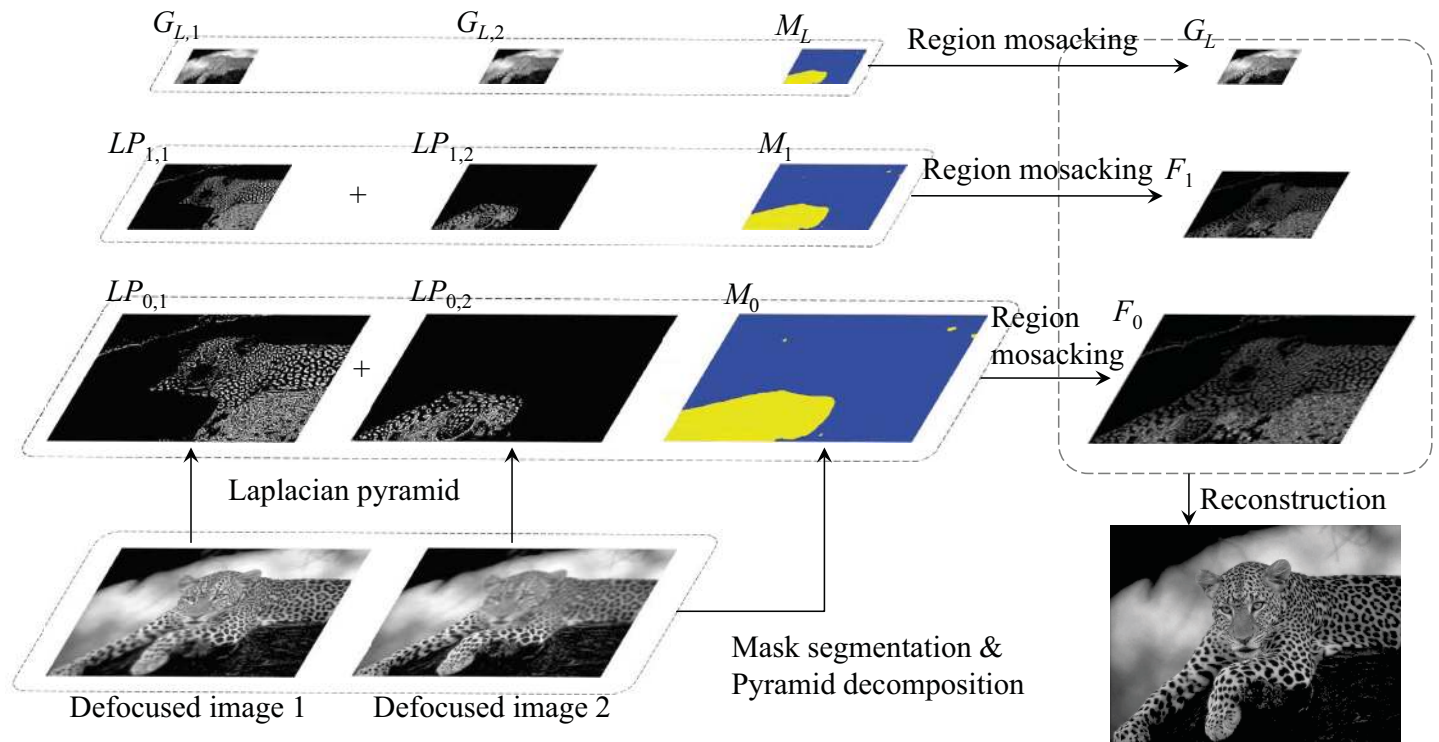


Fig 4. Illustration of RMLP with two multi-focus images and three pyramid layers.

<https://doi.org/10.1371/journal.pone.0191085.g004>

Region mosaicking on Laplacian pyramid

According to Eq 9, the pixel-level fusion pyramid values are selected according to the difference-of-Gaussian magnitudes of their source images. It is found that regions selected for fusion in different pyramid layers are not similar figures, and the pixels which need to be reconstructed often come from more than one multi-focus images. As a result, the original clear information is lost and distortion is introduced. Therefore, the region based Laplacian pyramid fusion scheme, i.e., RMLP is proposed as following.

In Fig 4, the focus label mask is also decomposed into the pyramid of three layers, $M_l, l = 0, 1, 2$. In M_l blue and yellow regions indicate respectively focused regions of the defocused image 1 and 2. $LP_{l,1}$ and $LP_{l,2}, (l = 0, 1)$ are the Laplacian pyramid of two defocused images, $G_{L,1}$ and $G_{L,2}$ are the base images of two Gaussian pyramids. F_0 and F_1 denote the fused Laplacian pyramids. The operator '+' denotes the fusion operation of RMLP.

In RMLP, the segmentation mask $M(i, j)$ is firstly decomposed into a mask pyramid, $M_k(i, j), k \in [0, K]$, which is then used to supervise a region level fusion of LP. The corresponding formulation is as follows,

$$F_k(i, j) = LP_{k, M_k(i, j)}(i, j), k \in [0, K]. \tag{10}$$

Where $F_k(i, j)$ denotes the fusion result of pixel (i, j) at the k^{th} layer, $M_k(i, j)$ denotes the mask label of the k^{th} layer LP, i.e., which multi-focus image should be selected for fusion at pixel (i, j) .

According to Eqs 8–10 the image fusion of the top pyramid is performed as

$$F_{K-1}(i, j) = G_{K-1, M_{K-1}(i, j)}(i, j). \tag{11}$$

Where $K - 1 = L = 2$ is the number of layers of LP, $F_{K-1}(i, j)$ denotes the fusion result of pixel (i, j) in the top layer image (K^{th} layer). $G_{K-1, M_{K-1}(i, j)}(i, j)$ denotes the value of (i, j) of the base image. $M_{K-1}(i, j) = \arg \max_n |SML_n^{K-1}(i, j)|, n \in [1, N]$ denotes the label of pixel (i, j) of K^{th} layer mask image, calculated by the DBRG segmentation algorithm in Section 3.

In the region based fusion procedure, M_{k+1} is the down-sampling copy of M_k , so corresponding regions in M_k and M_{k+1} are same figures. Consequently, most of the pixels (except for the pixels that lie on region transitive zones) in the fused image are reconstructed with pixel values from the same multi-focus images. As a result, much of the original clear information is kept and distortion is reduced.

When reconstructing the boundary areas, information from more than one multi-focus image is used. This may induce slight focus information loss in the transitive zones around the boundary areas. However, the usage of more than one multi-focus image information guarantees the gradual transformation of a transitive zone so that the block artifacts of the fused images is eliminated.

With the fusion results in the fused LPs ($F_k, k \in [0, K)$), the low-pass-filtered each layer of Gaussian pyramid (G_k), and the fused top layer image of LP ($F_{K-1}(i, j)$), a reconstruction procedure is then carried out with

$$\begin{cases} G_{K-1}(i, j) = LP_{K-1} \\ G_k(i, j) = LP_k + G_{k+1}^*, k \in [0, K - 1) \end{cases} \quad (12)$$

Where G_k and G^* are derived from Eqs 7 and 8. The reconstruction starts from the bottom layer Gaussian pyramid, G_0 ; then it iteratively calculates the G_{K-1} . According to the Eq 12, a Gaussian pyramid can be iteratively calculated from the top layer of the Laplacian pyramid. Finally, the bottom image of the Gaussian pyramid, G_0 , is precise reconstructed all-in-focus image.

Experiment results

A great number of the fusion approaches are carried in grayscale image. For preserving more information, however, the proposed RMLP is developed for color image fusion. RMLP is utilized to fuse the multi-focus images on R, G and B channels of RGB color image, respectively. Then an all-in-focus image is synthesized from the fused results of three channels. In this section, our proposed approach is evaluated and compared with the others in captured two datasets by a microscope.

Datasets

Two datasets (S1 and S2 Datasets) are collected including fifteen defocused images of the top of a bullet and twelve defocused images of a bee's body, respectively. The S1 Dataset is composed of image sequence that is captured at various depths by the microscope. The S2 Dataset is a sequence of synthetic images. Both S1 and S2 Datasets are used for subjective evaluation and objective evaluation respectively. As photographing equipment, the maximum magnification of the microscope of magnification 200 is used, which is made from the objective of the magnification 10 and the eyepiece of the magnification 20. The two datasets belong in the typical applications of forensic and biological fields.

Subject evaluation

Fig 5(a)–5(f) shows six samples of the first dataset, which are the image sequence under different focus parameters of the bullet. For the depth variation of the 'top of bullet' object, each image has only a stripe region in focus as indicated by a color stripe in the mask image (Fig 5(h)).

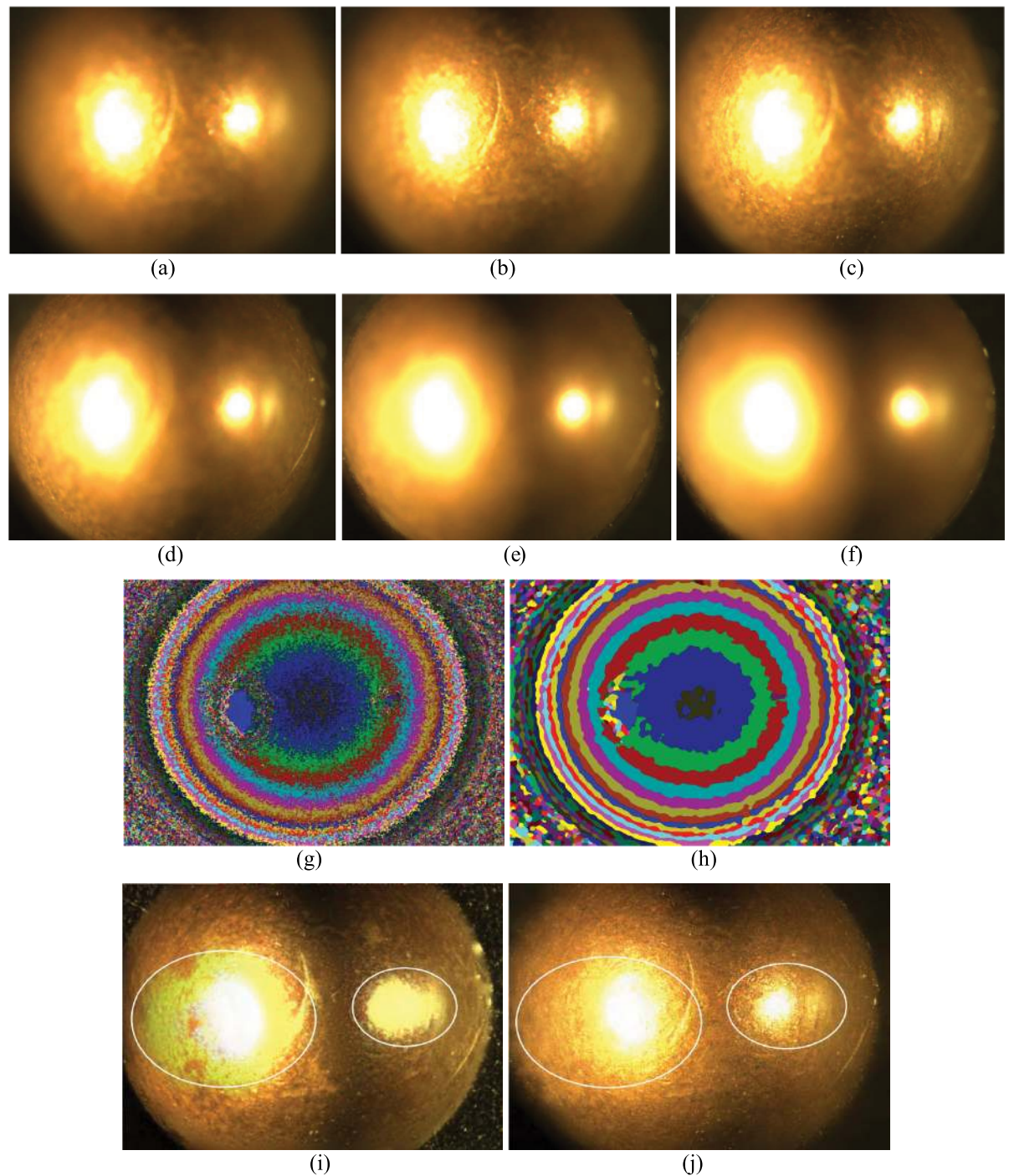


Fig 5. Illustration of the S1 Dataset and its fusion results. (a)-(f) are six random sampled examples from fifty multi-focus images, (g) is the mask image with only EOF measurement, (h) is the mask image with the proposed DBRG segmentation algorithm. (i) is the fusion result of the Laplacian pyramid (LP) method [11] and (j) is the fusion result of the proposed RMLP.

<https://doi.org/10.1371/journal.pone.0191085.g005>

When the pixel level focus measurement is used to calculate focus pixels, it can be seen in Fig 5(g) that the mask image is noisy and have many errors.

When the DBRG algorithm of a proper R (see Eq 5) parameter is employed, clear focus regions are segmented, as shown in Fig 5(h). In Fig 5(i) and 5(j) the fusion results of the

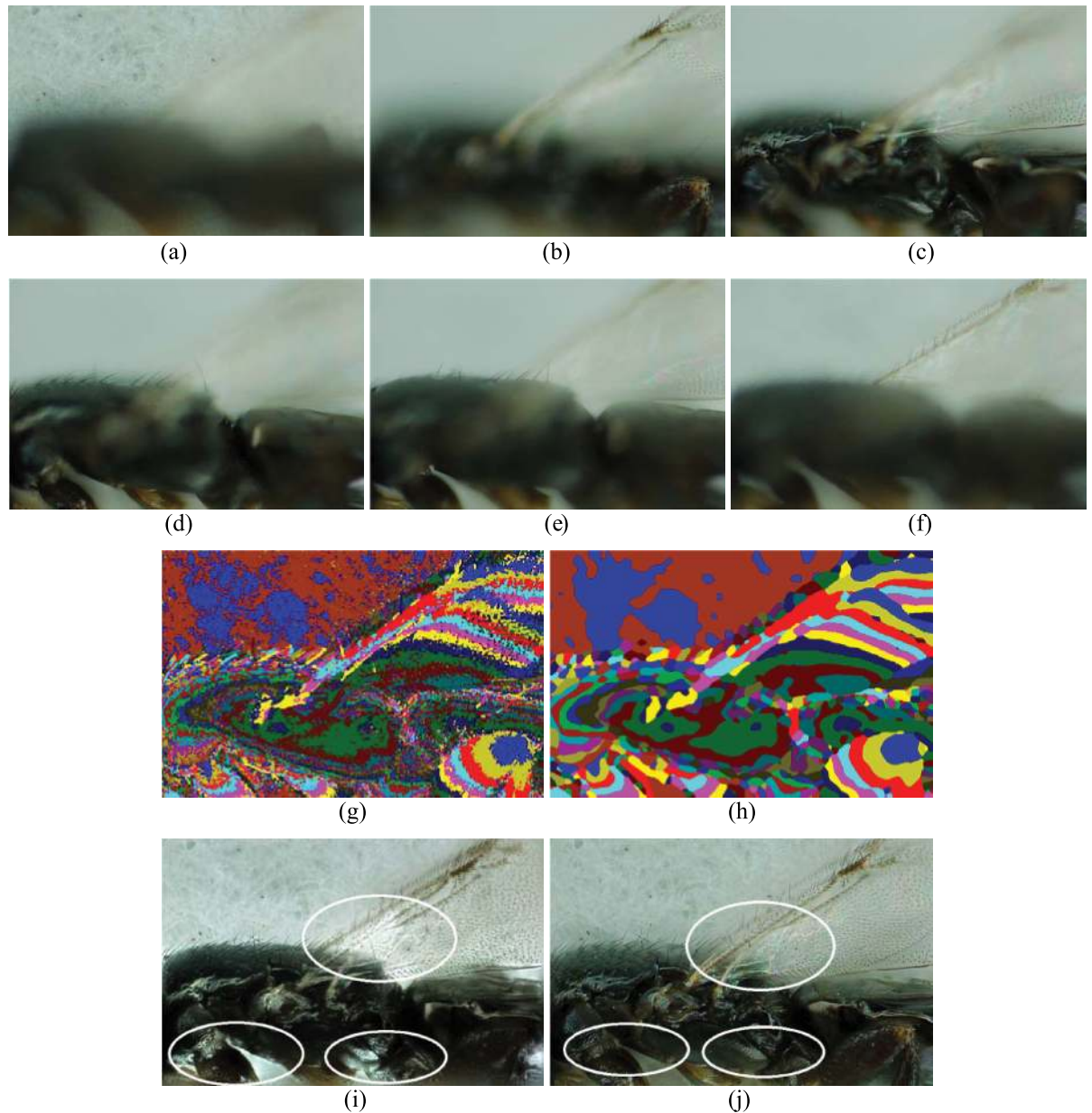


Fig 6. Illustration of the S2 Dataset and its fusion results. (a)-(f) are six random sampled examples from sixty multi-focus images, (g) is the mask image with only EOF measurement, (h) is the mask image with the proposed DBRG segmentation algorithm. (i) is the fusion result of the Laplacian pyramid (LP) method and (j) is the fusion result of the proposed RMLP.

<https://doi.org/10.1371/journal.pone.0191085.g006>

Laplacian Pyramid (LP) method [11] and the proposed RMLP approach are compared. It can be seen that in the highlighted regions with two circles in Fig 5(i) and 5(j), the fusion result of LP method has some color distortion, while the result of RMLP approach only has minimal distortion which validates the effectiveness of the proposed region mosaicing strategy in reducing distortion. Like Figs 5 and 6 shows also six samples of the second dataset with more complex texture feature and 3D shape. The results and comparison in Fig 6(i) and 6(j) demonstrate the performance of the RMLP approach over pixel level LP approach.

Objective evaluation

For objective evaluation, the images of the second dataset are fused and computed the difference between fused result and the ground-truth. The fusion precisions are evaluated by Root Mean Squared Error (RMSE) and Structural Similarity (SSIM) [43] [44] [45]. RMSE is defined as

$$RMSE = \left(\frac{\sum_{x=1}^X \sum_{y=1}^Y (I(x, y) - I'(x, y))^2}{X \cdot Y} \right)^{\frac{1}{2}} \tag{13}$$

Where X and Y are the width and height of the fused image, respectively. $I(x, y)$ denotes the value of the pixel (x, y) in the fused image and $I'(x, y)$ is the ground-truth of the pixel (x, y) as the reference value. RMSE is inversely proportional to the performance of the fusion algorithm. If a perfect ‘all-in-focus’ image is achieved, the RMSE will be close to zero.

SSIM can be calculated as:

$$SSIM = \frac{(2\mu_g\mu_f + C_1)(2\sigma_{gf} + C_2)}{(\mu_g^2 + \mu_f^2 + C_1)(\sigma_g^2 + \sigma_f^2 + C_2)} \tag{14}$$

Where μ_g denotes the mean intensity of all pixels in a sliding window of the ground-truth, μ_f denotes the mean intensity of all pixels in corresponding window of fused image, σ_g and σ_f are the statistical dispersions of all pixels the two windows, σ_{gf} is their covariance, $C_1 = (K_1 \times L)$ and $C_2 = (K_2 \times L)$ are two variables to stabilize the division with weak denominator. K_1 and K_2 are two constants, L is the dynamic range of the pixel values. In the experiments, K_1 , K_2 and L are set to 0.01, 0.03 and 255 [44]. As a similarity measure, SSIM compares fused image and the ground-truth. A larger value of SSIM indicates that the result is more consist with the ground-truth.

In Fig 7 two types of mask image, stripe and circle, are used to fuse multi-focus image. In Fig 7(a) and 7(b), the results of RMSE and SSIM are used to determine the value of parameter R . It can be seen in Fig 7(a) that when $R = 8 \sim 16$, the highest RMSE is obtained. It can be seen in Fig 7(b) that when $R = 8$, the largest SSIM is obtained. According to the above observations, $R \in [8, 16]$ is selected as a fine turned parameter. It should be mentioned that the parameters are selected for an image with 720×480 pixels. It has already been observed in experiments that the value of R should be scaled proportional to the image size. A larger value of R not only reduces the accuracy but also delay the fusion process.

Fig 7(c) and 7(d) show that RMSE and SSIM have the lowest and highest values respectively if the pyramid has 6 or 7 layers. More or less layer cannot achieve optimizations for the fused results. Therefore it is crucial that we must set the appropriate number of layers for the objective quality of the fused image.

In Fig 7(e) and 7(f), RMLP is compared with six state of the art methods that include pulse coupled neural network (PCNN) [1], nonsubsamped contourlet transform (NSCT) [21], interactive digital photomontage (IDP) [32], Laplacian pyramid (LP) [11], the contrast pyramid (CP) [16], Discrete Wavelet Transform (DWT) [46], the Shift-invariant Discrete Wavelet Transform (SIDWT) [25], Principal Component Analysis(PCA) [24] and sparse reconstruction method [29]. PCNN is implemented according to the descriptions in [1]. The programs of other approaches are downloaded from the authors’ websites [47] [48].

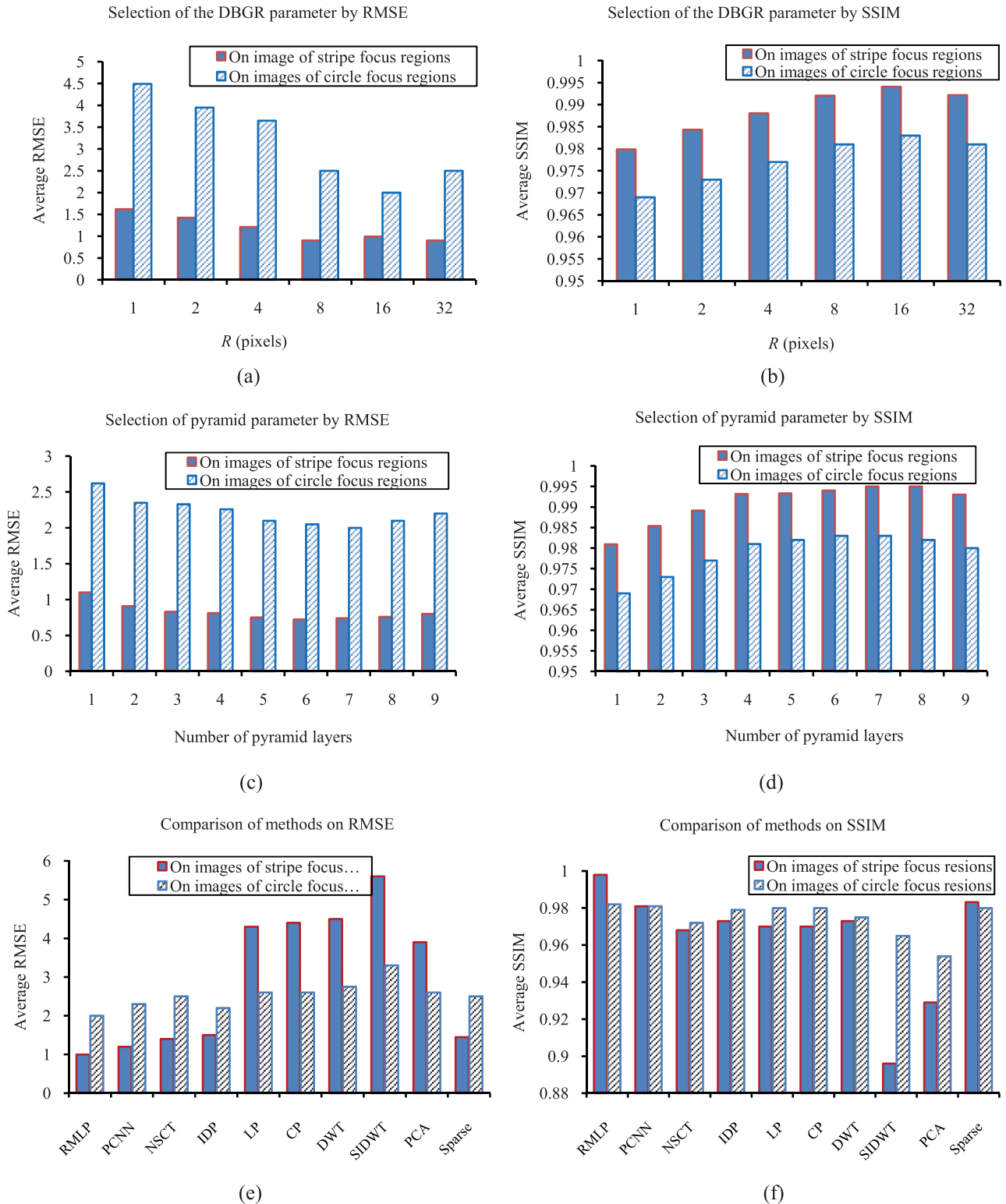


Fig 7. Illustration of quantitative evaluation (a) RMSE and (b) SSIM under different DBRG radius for the focus region mask segmentation. (c) RMSE and (d) SSIM with different pyramid layers. (e) and (f) are comparisons of nine methods.

<https://doi.org/10.1371/journal.pone.0191085.g007>

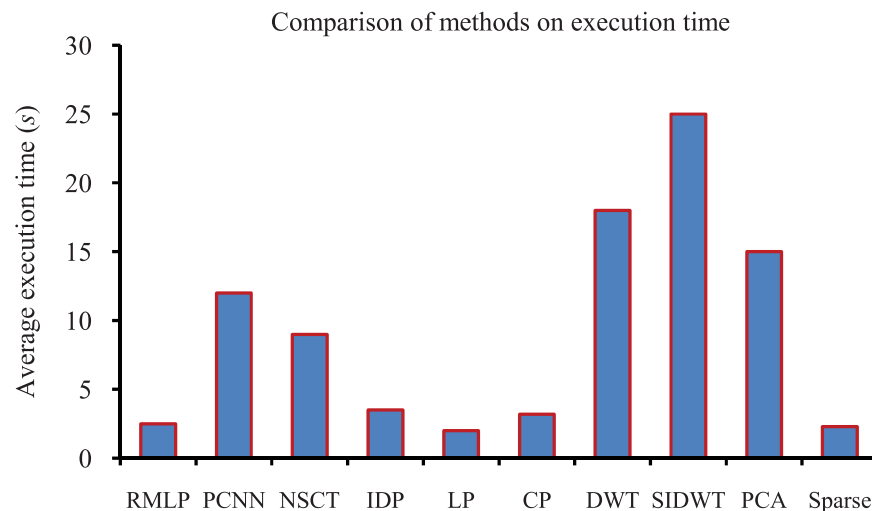


Fig 8. Average execution time of RMLP and other methods.

<https://doi.org/10.1371/journal.pone.0191085.g008>

In Fig 7(e), the RMSE of RMLP is the lowest comparing with other methods, which indicates that RMSE has highest precision for fused images. In Fig 7(f), both the fusion result of RMLP and ground-truth have highest similarity among seven approaches, which shows that RMLP retains the most information in the original image. To sum up, the proposed approach, RMLP, improves the performance of the state of the arts.

Considering the computation issue, the existing PCNN method has significantly higher computational complexity for iterative operations. By contrast, our RMLP approach is significantly simpler and more efficient. The average execution time is also compared, which is shown in Fig 8. The average execution time of PCNN approach is about 12.0 seconds to fuse an image of 720×480 pixels on an Intel Core i3-2130CPU of 3.4 GHz. The proposed approach, RMLP, however, spends just 1.35 seconds. In contrast, the IDP approach that uses a graph-cut optimal algorithm to calculate focus region mask spends 2.6 seconds on average.

Conclusions

Pyramid based image fusion approaches, such as Wavelet or Contrast Pyramids, are classic methods for image fusion. When these methods are applied to multi-focus image fusion, the region continuity characteristics are not fully explored. In this paper we propose a new multi-focus image fusion approach, called RMLP, which explores unique characteristics of multi-focus imaging and transfers the classical pixel-based pyramid fusion to region-based. The performance of RMLP is evaluated in both objective and subjective datasets. The experiment results show that RMLP has lowest error and best robustness comparing with the existing methods.

A well-known problem of existing popular approaches is that all multi-focus images need to be aligned with each other. The miss-alignment of images may cause failure on the fusion results. In order to insure alignment, the multi-focus images of our data set were captured by a stable platform (a desktop microscope).

Supporting information

S1 Dataset. Image sequence captured at various depths by the microscope.
(ZIP)

S2 Dataset. Image sequence of synthetic images.
(ZIP)

Acknowledgments

This work is supported by the National Science Foundation of China (Grant No. 61501132, Grant No. 61370084), the China Postdoctoral Science Foundation No. 2016M591515, the Heilongjiang Postdoctoral Sustentation Fund with No. LBH-Z14055, Harbin Application Technology Research and Development Project (Grant No. 2016RAQXJ063, Grant No. 2016RAXXJ013).

Author Contributions

Conceptualization: Jianguo Sun.

Data curation: Liang Kou.

Formal analysis: Jianguo Sun.

Investigation: Qilong Han.

Methodology: Liang Kou, Liguozhang, Qilong Han.

Resources: Kejia Zhang.

Software: Liguozhang, Zilong Jin.

Writing – original draft: Kejia Zhang.

Writing – review & editing: Kejia Zhang.

References

1. Wang ZB, Ma Yd, Gu J. Multi-focus image fusion using PCNN. *Pattern Recognition*. 2010 Dec; 43(6): 2003–2016. <https://doi.org/10.1016/j.patcog.2010.01.011>
2. Johnson J L, Padgett M L. PCNN models and applications. *IEEE Transactions on Neural Networks*. 1999; 10(3):480. <https://doi.org/10.1109/72.761706> PMID: 18252547
3. Favaro P, Soatto S, Burger M, Osher SJ. Shape from defocus via diffusion. *IEEE transactions on pattern analysis and machine intelligence*. 2008; 30(3): 518–531. <https://doi.org/10.1109/TPAMI.2007.1175> PMID: 18195444
4. Zhuo S, Sim T. Defocus map estimation from a single image. *Pattern Recognition*. 2017; 44(9):1852–1858. <https://doi.org/10.1016/j.patcog.2011.03.009>
5. Kausar N, Majid A, Javed S G. Developing learning based intelligent fusion for deblurring confocal microscopic images. *Engineering Applications of Artificial Intelligence*. 2016; 55:339–352. <https://doi.org/10.1016/j.engappai.2016.08.006>
6. Dogan H, Ayas S, Ekinici M. Auto-focusing with multi focus color image fusion based on curvelet transform on microscopic imaging. *IEEE International Conference on Electrical and Electronics Engineering*. 2016; 253–258.
7. Tsai DC, Chen HH. Reciprocal Focus Profile. *IEEE Transactions on Image Processing*. 2012; 21(2):459–468. <https://doi.org/10.1109/TIP.2011.2164417> PMID: 21843991
8. Ma Z, Wen J, Zhang C, Liu Q, Yan D. An effective fusion defogging approach for single sea fog image. *Neurocomputing*. 2016; 173: 1257–1267. <https://doi.org/10.1016/j.neucom.2015.08.084>
9. Hariharan R, Rajagopalan AN. Shape-from-focus by tensor voting. *IEEE Transactions on Image Processing*. 2012; 21(7):3323–3328. <https://doi.org/10.1109/TIP.2012.2190612> PMID: 22434802
10. Hu Y, Sirlantzis K, Howells G. Signal-Level Information Fusion for Less Constrained Iris Recognition Using Sparse-Error Low Rank Matrix Factorization. *IEEE Transactions on Information Forensics and Security*. 2016; 11(7):1549–1564. <https://doi.org/10.1109/TIFS.2016.2541612>
11. Wang W, Chang F. A Multi-focus Image Fusion Method Based on Laplacian Pyramid. *JCP*. 2011; 6(12):2559–2566.

12. Pradnya P M, Sachin D R. Wavelet based image fusion techniques. IEEE International Conference on Intelligent Systems and Signal Processing. 2013; 77–81.
13. Tian J, Chen L. Multi-focus image fusion using wavelet-domain statistics. IEEE International Conference on Image Processing. 2010 Sep; 1205–1208.
14. Petrovic VS, Xydeas CS. Gradient-based multiresolution image fusion. IEEE Transactions on Image Processing. 2004; 13(2):228–237. <https://doi.org/10.1109/TIP.2004.823821> PMID: 15376943
15. Zhou Z, Li S, Wang B. Multi-scale weighted gradient-based fusion for multi-focus images. Information Fusion. 2014; 20(1):60–72. <https://doi.org/10.1016/j.inffus.2013.11.005>
16. Dong Y B, Li M J, Li J. Image Fusion Algorithm Based on Contrast Pyramid and its Performance Evaluation. Applied Mechanics & Materials. 2014; 525: 711–714. <https://doi.org/10.4028/www.scientific.net/AMM.525.711>
17. Toet A. Image fusion by a ratio of low-pass pyramid. Pattern Recognition Letters. 1989; 9(4):245–253. [https://doi.org/10.1016/0167-8655\(89\)90003-2](https://doi.org/10.1016/0167-8655(89)90003-2)
18. Rockinger Oliver Image Sequence Fusion Using a Shift-Invariant Wavelet Transform. IEEE International Conference on Image Processing. 1997; 3: 288–291.
19. Lewis JJ, Callaghan RJO, Nikolov SG, Bull DR, Canagarajah N. Pixel- and region-based image fusion with complex wavelets. Information fusion, 2007; 8(2):119–130. <https://doi.org/10.1016/j.inffus.2005.09.006>
20. Chabi N, Yazdi M, Entezarmahdi M. An efficient image fusion method based on dual tree complex wavelet transform. IEEE Machine Vision and Image Processing. 2014; 403–407.
21. Yang Y., Tong S., Huang S., Lin P. Multifocus Image Fusion Based on NSCT and Focused Area Detection. IEEE Sensors Journal. 2015; 15(5):2824–2838.
22. Yang H, Jin X, Zhou D. Multi-focus Image Fusion Using Guided Filtering and Complex Contourlet Transform. Journal of Signal Processing. 2016.
23. Srivastava R, Khare A. Multifocus noisy image fusion using contourlet transform. International Conference on Contemporary Computing. 2015; 497–502.
24. Jing H, Vladimirova T. Novel PCA based pixel-level multi-focus image fusion algorithm. IEEE Adaptive Hardware and Systems. 2014; 135–142.
25. Saha A, Bhatnagar G, Wu Q. Mutual spectral residual approach for multifocus image fusion. Digital Signal Processing. 2013; 23(4):1121–1135.
26. Zhang Y, Ge L. Efficient fusion scheme for multi-focus images by using blurring measure. Digital signal processing, 2009; 19(2):186–193. <https://doi.org/10.1016/j.dsp.2008.11.002>
27. Yang B, Li ST. Multifocus Image Fusion and Restoration With Sparse Representation. IEEE Transactions on Instrumentation & Measurement. 2010; 59(4):884–892. <https://doi.org/10.1109/TIM.2009.2026612>
28. Li H, Wei S, Chai Y. Multifocus image fusion scheme based on feature contrast in the lifting stationary wavelet domain.
29. Yin H, Li Y, Chai Y, Liu Z, Zhu Z. A novel sparse-representation-based multi-focus image fusion approach. Neurocomputing. 2016; 216(C):216–229. <https://doi.org/10.1016/j.neucom.2016.07.039>
30. Bai X, Zhang Y, Zhou F, Xue B. Quadtree-based multi-focus image fusion using a weighted focus-measure. Information Fusion. 2015; 22: 105–118. <https://doi.org/10.1016/j.inffus.2014.05.003>
31. Kubota A, Aizawa K. Reconstructing arbitrarily focused images from two differently focused images using linear filters. IEEE Transactions on Image Processing. 2005; 14(11):1848–1859. <https://doi.org/10.1109/TIP.2005.854468> PMID: 16279184
32. Agarwala A, Dontcheva M, Agrawala M, Drucker S, Colburn A, Curless B, Salesin D, Cohen M. Interactive digital photomontage. ACM Transactions on Graphics. 2004; 23(3):294–302. <https://doi.org/10.1145/1015706.1015718>
33. Eckhorn R, Reitboeck H, Arndt M, Dicke P. Feature linking via synchronization among distributed assemblies: simulation of results from cat cortex. Neural computation, 1990; 2(3):293–307. <https://doi.org/10.1162/neco.1990.2.3.293>
34. Gayathri N, Deepa P L. Multi-focus color image fusion using NSCT and PCNN. IEEE International Conference on Communication Systems and Networks. 2017; 173–178.
35. Yang Hengfen and Jin Xin and Zhou Dongming. Block medical image fusion based on adaptive PCNN. IEEE International Conference on Software Engineering and Service Science. 2015;330–333.
36. Chen Y, Qin Z PCNN-Based Image Fusion in Compressed Domain. Mathematical Problems in Engineering. 2015; 2015(6):1–9.

37. Hariharan H, Koschan A, Abidi M. An Adaptive Focal Connectivity Algorithm for Multifocus Fusion. *IEEE Conference on Computer Vision and Pattern Recognition*. 2007 Jun; 124(1):1-6.
38. Huang W, Jing ZL. Evaluation of focus measures in multi-focus image fusion. *Neural computation*, 1990; 2(3):293–307.
39. Nayar Nakagawa. Shape from focus. *IEEE transactions on pattern analysis and machine intelligence*, 1994; 16(8):824–831. <https://doi.org/10.1109/34.308479>
40. Zheng YH, Byeungwoo J, Xu DH, Wu QM, Zhang H. Image segmentation by generalized hierarchical fuzzy C-means algorithm. *Journal of Intelligent and Fuzzy Systems*. 2015; 28(2):4024–4028.
41. Li J, Li XL, Yang B, Sun XM. Segmentation-Based Image Copy-Move Forgery Detection Scheme. *IEEE Transactions on Information Forensics & Security*. 2015; 10(3):507–518. <https://doi.org/10.1109/TIFS.2014.2381872>
42. Shang Y, Bouffanais R. Influence of the number of topologically interacting neighbors on swarm dynamics. *Scientific Reports*. 2014; 4(8):4184. <https://doi.org/10.1038/srep04184> PMID: 24567077
43. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*. 2004; 13(4):600–612. <https://doi.org/10.1109/TIP.2003.819861> PMID: 15376593
44. Goshtasby AA, Nikolov S. Image fusion: Advances in the state of the art. *Information Fusion*. 2007; 8(2):114–118. <https://doi.org/10.1016/j.inffus.2006.04.001>
45. Li S, Kang X, Fang L, Hu J, Yin H. Pixel-level image fusion: A survey of the state of the art. *Information Fusion*. 2017; 33: 100–112. <https://doi.org/10.1016/j.inffus.2016.05.004>
46. Pajares G, Cruz JM. A wavelet-based image fusion tutorial. *Pattern Recognition*. 2004; 37(9):1855–1872. <https://doi.org/10.1016/j.patcog.2004.03.010>
47. <http://grail.cs.washington.edu/projects/photomontage/>.
48. <http://www.metapix.de/toolbox.htm/>.