



Template Mask Based Image Fusion Built-in Algorithm for Wide Field Fundus Cameras

Jun Wu¹(✉), Mingxin He¹, Yang Liu¹, Hanwen Zhang¹, Jianchun Zhao²,
Xiaohou Shen², Jiankun Liu², Gang Yang³, Xirong Li³, and Dayong Ding²

¹ School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China

junwu@nwpu.edu.cn, {hemingxin, 2021202082, zhanghanwen}@mail.nwpu.edu.cn

² Vistel AI Lab, Visionary Intelligence Ltd., Beijing 100080, China

{jianchun.zhao, xiaohou.shen, jiankun.liu, dayong.ding}@vistel.cn

³ Key Lab of DEKE, Renmin University of China, Beijing 100872, China

{yanggang, xirong}@ruc.edu.cn

Abstract. When a typical wide field fundus camera takes two fundus images as a pair, two patterns of illumination beams (the top and bottom, left and right) are turned on respectively. Due to the influence of the illumination beams, the reflected and scattered light haze are observed in neighboring regions, which results in partial or even complete occlusion of retinal structures. In this paper, we propose a novel template mask based image fusion algorithm for wide field fundus cameras, which splices available higher-quality regions from these two images in one pair, and fuses them into a single high-quality image, which tries to keep complete retinal structure. First, the region of interest (ROI) are obtained based on the Hough circle transform, and these two images are adjusted by color and brightness normalization based on Poisson fusion. Then, a customized template mask is designed to fuse the higher-quality regions from this image pair. Finally, image enhancement in brightness is carried out to further improve final image quality. Experimental results on a variety of wide field fundus image pairs demonstrate that the proposed method is effective, our method increases information entropy by 0.051, standard deviation by 1.726, average gradient by 0.148 and spatial frequency by 0.233.

Keywords: Wide field fundus camera · Image processing · Image fusion

1 Introduction

In typical wide field fundus cameras, there will be a strong stray light effect (fog-like phenomenon) close to the built-in light sources, resulting in occlusion and irreversible interference to retinal structures or lesions, so that the wide field fundus camera cannot shoot only once to achieve high-quality imaging. A typical solution is to divide the

four illumination beams around the observation axis into two lighting patterns (top and bottom, left and right), and turn them on alternately to capture two wide field fundus images as a capturing pair, as shown in Fig. 1. Therefore, it is necessary to develop an image fusion algorithm for wide field fundus cameras to obtain a high-quality fundus image with relatively complete retinal structures.

In traditional image fusion problems, the corresponding pixel pairs of two input images are generally aligned (after image registration) and complementary. However, wide field fundus camera images belong to a different case. High-quality central areas of the image pair can be both retained and complementary. However, low-quality areas close to light sources are almost unusable, only the relatively high-quality area in the other image can be reserved. From this viewpoint, they are choose-one-from-two as the same image area is in different light states in two patterns. In addition, the global and local brightness of two input images are significantly different due to two different lighting patterns. As a result, the boundary effects between reserved and discarded regions are prone to produce obvious visual differences, resulting in unsatisfactory overall image fusion results.

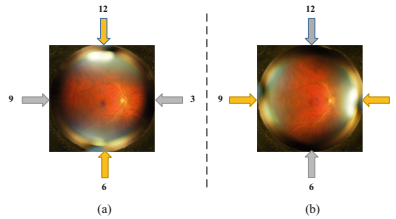


Fig. 1. A pair of wide field fundus images. (a) The first image (top and bottom light sources at 6 and 12 o'clock). (b) The second image (left and right light sources at 3 and 9 o'clock)

In this paper, we propose a template mask based image fusion algorithm for wide field fundus cameras to obtain high-quality wide field fundus images. It was observed that two wide field fundus images in the same pair are spatially complementary. That is to say, there is a low-quality region near the light source in one image, a corresponding better-quality region occurs in the other image at the same position. To take full advantage of this property, an image fusion algorithm based on template mask is proposed, which splices the available parts of two wide field fundus images and fuses them into a relatively complete and high-quality fundus image.

2 Related Work

To preserve original accurate information as much as possible, pixel-wise image fusion algorithms can be applied [1, 2], which can be divided into two categories: image fusion based on spatial domain and image fusion based on transform domain.

Image fusion based on spatial domain is generally performed directly on the gray space of image pixels. A most straightforward method is to use the maximum method or the weighted average method [4] to select the pixel from two input images, which operates

directly on the target pixel without considering the correlation between neighboring pixels. When there is large information complementarity between images to be fused, a region-based image fusion method can be applied, the fusion coefficients of different to-be-used images can be determined according to the feature relations between the pixels in a rectangular window at a certain position. Zhang [5] obtained the Laplacian energy of the input image to measure the focusing degree, and realized the image fusion by a sliding window. However, the computational complexity of this algorithm is high. Principal component analysis is also a typical spatial domain method, Zhu [6] searched for principal components of images by dimension reduction, and determined the weight of each fusion image according to the energy of principal components. Chen [3] proposed an image fusion algorithm based on edge detection. The improved ROEWA (Ratio of Exponentially Weighted Averages) operator is used to detect image edge. Different image fusion rules are set according to the high-frequency region and low-frequency region. In addition, image fusion methods based on spatial domain also include false color image fusion [7], image fusion based on modulation [8], image fusion based on statistics [9], and so on.

Common fundus image fusion methods based on transform domain mainly include fusion based on pyramid transform [10–13, 24] and wavelet transform [14–18, 21, 23]. The image fusion based on pyramid transform extracts the image detail information on different decomposition scales and has a good fusion effect. However, after pyramid decomposition, the data between decomposition layers is redundant, and the high-frequency information might be seriously lost. The image fusion based on wavelet transform can not only extract low-frequency information, but also obtain high-frequency detailed information. However, because the wavelet transform uses row and column down sampling, the image is not translation invariant, which easily leads to the distortion of the fused image.

Different from traditional image fusion, image fusion for wide field fundus images should not only consider correlation between pixels, but also abandon low-quality regions where retinal structures near light source is covered. Paul et al. [22] proposed an image fusion algorithm, where a mask generated by spectral analysis is used to score the visibility of each pixel from source image, and each pixel in output image takes the corresponding source pixel with the highest score. Its mask derived bigger transmission region, unfriendly to strong fog-effect images. In this paper, based on the complementarity between two images from the same pair, an image fusion algorithm based on well-defined template mask is proposed.

3 The Proposed Method

The proposed image fusion algorithm for wide field fundus cameras is shown in Fig. 2. The algorithm consists of four parts: wide field fundus image pre-processing, color and brightness normalization, image fusion based on template mask, and adaptive brightness adjustment. For the input wide field fundus image pair, pre-processing is applied first to improve image quality. Then, the color and brightness of the two pre-processed images are normalized by Poisson fusion to reduce color and brightness differences. Next, the high-quality regions of the two images are selected for fusion based on a template mask.

Finally, the fusion image with low brightness is enhanced to further improve the overall image quality.

3.1 Wide Field Fundus Image Pre-processing

Image Defogging and Registration. Due to the imaging characteristics of wide field fundus cameras, the reflected and scattered light haze are observed in images, and there is a difference of capturing times between the two wide field fundus images, leading to possible offsets. Therefore, the image should be defogged and registered first. In this paper, dark channel prior-based defogging algorithm [19] is used to obtain a wide field fundus image with better quality to see clearer retinal structures. An image registration algorithm based on SIFT feature points is used. Firstly, the brightness and contrast of the image are improved to highlight the retinal detail. The SIFT feature points are detected, which are filtered by RANSAC, and then the input images are registered. Through registration, the center consistency of the two images can be guaranteed, the pixel error at the joint after image fusion can be effectively reduced, and the stitching accuracy can be improved.

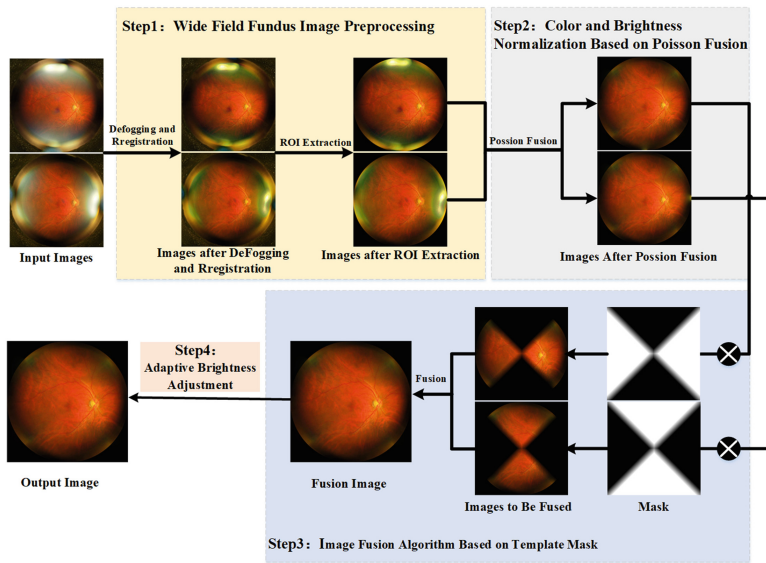


Fig. 2. Block diagram of image fusion algorithm for wide field fundus cameras.

Region of Interest (ROI) Extraction. The effective retinal area of wide field fundus image is approximately circular. In order to avoid the interference of useless areas outside the retina, the circular retinal field of view (FOV) is extracted through Hough circle detection. To ensure that the two images remain registered, they must use the same circle to extract the FOV, so the average of the detected center and the average ROI size of the two images is taken as the final result, and ROI of the two images is extracted

with this result. As shown in Fig. 3, the retinal structures of the fundus become visible and they are of better quality after pre-processing.

3.2 Color and Brightness Normalization Based on Poisson Fusion

Different light sources are used in the shooting of two wide field fundus images in the same pair, which easily results in differences in brightness and color between these two images, and obvious boundary is traced in the fusion result. In order to reduce the influence of color and brightness differences on the fusion result, it is necessary to normalize the color brightness of these two images.

During the Poisson fusion process, these two images are considered as foreground and background respectively. It adjusts the color of the foreground image to that of the background image, effectively reducing the brightness and color difference between two images, weakening the image splicing boundary effect, and ensuring the overall color balance of the fusion result.

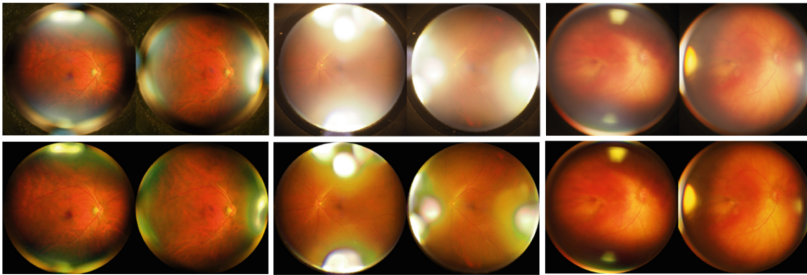


Fig. 3. Results before (the first row) and after (the second row) pre-processing.

3.3 Image Fusion Based on Template Mask

Two fundus images of the same subject have spatial complementarity. In the low-quality regions that must be discarded close to the light source, the image quality is relatively better at the corresponding position of the other image. The images are divided using a template mask according to the distribution characteristics of stray light. An image fusion algorithm based on a diagonal mask is proposed. The high-quality regions of the original image are selected for fusion based on this template mask.

As shown in Fig. 4, according to the presence of stray light, the wide field fundus images along the diagonal are divided into four parts. The top and bottom regions of Image A (A_1 , A_2) and the left and right areas of Image B (B_3 , B_4) are the area without stray light, which can be retained. Other areas are abandoned due to strong stray light.

Figure 4 (c) shows the corresponding relationship between the fusion result and the two input images in each region. If the mask is directly generated according to the above method for fusion, there will be obvious boundaries on the diagonal of the fused image.

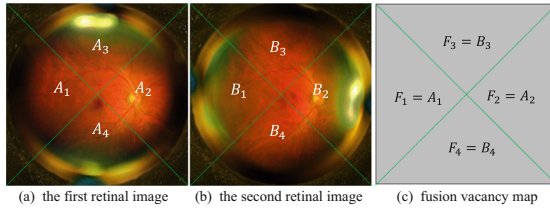


Fig. 4. Region division method of wide field fundus image and fusion image

To further improve image quality, the width w is expanded outward along the direction perpendicular to the boundary line of the image to be fused, so as to ensure that there are overlapping regions between two images to be fused. The weighted average method is applied to adjust the weight of pixel to eliminate the boundary effect. As shown in Fig. 5, for the diagonal overlapping region S_1 , reduce the weight of Image A and increase the weight of Image B along the arrow direction to make the boundary transition be smooth. Similarly, for S_2, S_3 and S_4 , the same method is employed realize smooth transition of color at the mask boundary.

The masks are generated according to the above method, and input image is multiplied by the corresponding masks to select the valid region. Finally, these images are combined to obtain fused image with complete retinal structures.

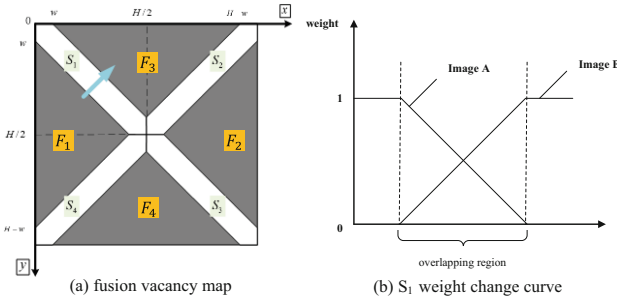


Fig. 5. Set the overlapping regions. (a) fusion vacancy map; (b) the weight change curve of the overlapping region S_1 . Along the arrow direction, the weight of Image A in the fusion image decreases from 1 to 0, and that of Image B in the fusion image increases from 0 to 1;

3.4 Adaptive Brightness Adjustment

Generally, the brightness of wide field fundus images is relatively low, which requires a suitable image enhancement step. To avoid the influence of low-quality area, the average pixel value in center region of Y channel in YUV color space is defined as the brightness. The average brightness of the fused image from the training set data is taken as the threshold. First, the gray values of all pixels in the 1% and 99% quartile are used as the pixel minimum P_{min} and maximum P_{max} respectively. Then, the pixel values greater

than P_{max} and less than P_{min} are truncated as P_{max} or P_{min} . Finally, the image is stretched to 0–255 to obtain the final image enhancement result.

4 Evaluations

4.1 Data Set and Experimental Setup

The dataset used in our evaluation is a private dataset captured using Retivue with Olympus Air A01, a Portable Wide Field Fundus Camera. It consists of 56 pairs of image in 1920×1920 pixels, and each pair contains two wide field fundus images of the same subject, the shooting time interval between the two images is about 100 ms. The dataset is divided into two subsets: 26 image pairs as a training set (to determine the threshold in module 4) and 30 image pairs as a test set.

It is conducted in Intel Core I5 under Win10 based on Python 3.7 and Opencv3.4.2. The image fusion algorithm based on template mask has an adjustable parameter: overlapping area width w . If w is too small, there will still has boundary effect in the fusion result. If w is too large, the stray light region in the original images that should have been abandoned will still exist in the fusion result. Based on our experimental verification, $w = 100$ is the most appropriate value for the target data set.

4.2 Results and Analysis

The proposed image fusion algorithm for wide field fundus cameras includes four modules: module 1 - wide field fundus image pre-processing, module 2 - image color and brightness normalization based on Poisson fusion, module 3 - image fusion based on template mask, and module 4 - adaptive brightness adjustment.

Validation of Each Module. First, the effectiveness of the image fusion algorithm based on template mask (module 1 + module 3) is verified. Then, taking this as the baseline, the effectiveness of module 2 and module 4 are verified.

Table 1. Results of this algorithm and traditional image fusion algorithm.

Fusion Algorithm	Information Entropy	Standard Deviation	Average Gradient	Spatial Frequency
Image Fusion Algorithm Based on Pixel Definition [21]	6.017	38.050	0.523	1.429
Image Fusion Algorithm Based on Diagonal Mask Pairs(baseline)	6.068	39.776	0.671	1.662

According to Table 1, compared to the image fusion algorithm based on Pixel Definition [21], our proposed method is more effective, which increases information entropy by 0.051, standard deviation by 1.726, average gradient by 0.148 and spatial frequency by 0.233. Further, the module 2 and module 4 are successively added to the baseline. Results are shown in Table 2. All evaluation indexes increase, indicating that the image color and brightness normalization and adaptive brightness adjustment based on Poisson fusion can effectively improve the quality of fusion results.

Table 2. Results of each experiment scheme.

Experiment Scheme	Brightness	Information Entropy	Standard Deviation	Average Gradient	Spatial Frequency
module 1 + module 3	-	6.068	39.776	0.671	1.662
module 1 + module 2 + module 3	101	6.214	40.459	0.684	1.671
module 1 + module 2 + module 3 + module 4	110	6.328	43.773	0.747	1.832

Validation of the Overall Algorithm. In order to verify the effectiveness of the algorithm, three domain experts scores the fusion results from three aspects of image authenticity, image clarity and overall quality. The image authenticity is based on whether there is obvious noise or color distortion in images. Image clarity is according to whether images are clearly visible, and whether the details are legible, blurred, or even lost. Overall quality is combined with general features (such as brightness and contrast) and structural features (such as vascular clarity, vascular density, and macular area contrast) to comprehensively evaluate images. Since our method is aiming at a specific wide-field fundus camera and there is no common overlap area in images, it is difficult to directly compare with existing image fusion methods, so we mainly compare to a realted work [22] in a similar strategy.

We compare them with Paul et al. [22] to evaluate the effectiveness of the algorithm. According to Table 3, The proposed method is slightly lower than that in Paul et al. [22] in terms of image authenticity, because there are still obvious traces in the diagonal transition area of some images. But the image clarity and overall image quality are both higher than those in Paul et al. [22]. It can be seen that our proposed method is better and more effective. Some visual examples are provided in Fig. 6.

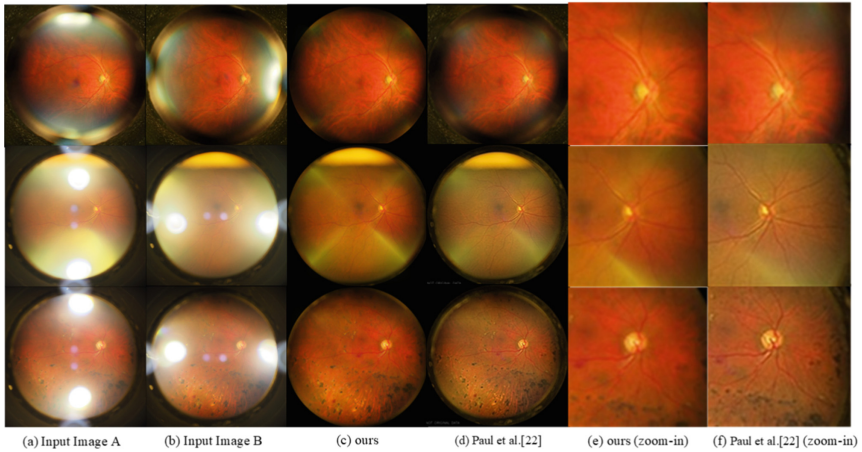


Fig. 6. Examples of original images and the comparison between our method and Paul's [22].

Table 3. Expert evaluation results (average of 30 pairs of data).

	Image clarity (0–10)	Image authenticity (0–10)	Overall quality (0–10)	Total (0–30)
Paul et al. [22]	7.944	7.756	7.676	23.356
Ours	8.144	7.589	7.711	23.444

5 Conclusions

In this paper, an image fusion algorithm for wide field fundus cameras is proposed. Image pre-processing is first conducted to eliminate the interference of invalid areas. An image fusion algorithm based on template mask is applied, which selects high-quality regions of wide field fundus images, and reduces the boundary effect of image fusion by weighted average method. Aiming at low brightness of wide field fundus images, a brightness adaptive image enhancement step is used to improve the information entropy of fused images and finally output a clear and high-quality wide field fundus image. Experimental results show that the proposed algorithm for wide field fundus cameras is effective and it has a better fusion results.

Acknowledgements. This work is supported by Natural Science Basic Research Plan in Shaanxi Province of China (2020JM-129), and BJNSF (No. 4202033).

References

1. Yang, B., Jing, Z.L., Zhao, H.T.: Review of pixel-level image fusion. *J. Shanghai Jiaotong Univ. (Sci.)* **15**(1), 6–12 (2010). <https://doi.org/10.1007/s12204-010-7186-y>

2. Singh, S., Anand, R.S.: Ripplet domain fusion approach for CT and MR medical image information. *Biomed. Signal Process. Control* **46**, 281–292 (2018)
3. Chen, J., Chen, L., Shabaz, M.: Image fusion algorithm at pixel level based on edge detection. *J. Healthc. Eng.* **7**, 1–10 (2021)
4. Min, X.: Application of improved weighted fusion in eliminating image stitching ghosting. *Softw. Eng.* **20**(5), 27–29 (2007)
5. Zhan, K., Teng, J., Li, Q., et al.: A novel explicit multi-focus image fusion method. *J. Inf. Hiding Multimedia Signal Process.* **6**(3), 600–612 (2015)
6. Zhu, Y. L., Liu, R. L., et al.: Research on medical image fusion based on PCA. In: 2018 International Conference on Computational, Modeling, Simulation and Mathematical Statistics, vol. 4 (2018)
7. Toet, A., Walraven, J.: New false color mapping for image fusion. *Opt. Eng.* **35**(3), 650–658 (1996)
8. Lv, X., Zhang, W.: Contrast modulation fusion method applied to medical images. *Comput. Technol. Dev.* **14**(012), 16–15 (2004)
9. Xia, Y., Leung, H., Bosse, E.: Neural data fusion algorithms based on a linearly constrained least square method. *IEEE Trans. Neural Netw.* **13**(2), 320–329 (2002)
10. Burt, P.J., Adelson, E.H.: The laplacian pyramid as a compact image Code. *IEEE Trans. Commun.* **31**(4), 532–540 (1983)
11. Toet, A., Ruyven, J.J.V., Valeton, J.M.: Merging thermal and visual images by a contrast pyramid. *Opt. Eng.* **28**(7), 789–792 (1989)
12. Toet, A.: A morphological pyramidal image decomposition. *Pattern Recogn. Lett.* **9**(4), 255–261 (1989)
13. Burt, P.J.: A gradient pyramid basis for pattern-selective image fusion. In: Proceedings of the Society for Information Display Conference. SID Press, San Jose, USA, pp. 467–470 (1992)
14. Li, H., Manjunath, B.S., Mitra, S.K.: Multi-sensor image fusion using the wavelet transform. *Graph. Models Image Process.* **57**(3), 235–245 (1995)
15. Chipman, L.J., Orr, T.M., Graham, L.N.: Wavelets and image fusion. In: Proceedings International Conference on Image Processing, pp. 248–251 (1995)
16. Guixi, L., Wanhai, Y.: Image fusion method and performance evaluation based on wavelet decomposition. *Acta Automatica Sin.* **28**(6), 927–934 (2002)
17. Wang, H.H.: A new multiwavelet-based approach to image fusion. *J. Math. Imaging Vis.* **21**(2), 177–192 (2004). <https://doi.org/10.1023/B:JMIV.0000035181.00093.e3>
18. Liu, D., Yang, F., Wei, H., et al.: Remote sensing image fusion method based on discrete wavelet and multiscale morphological transform in the IHS color space. *J. Appl. Remote Sens.* **14**(1), 016518 (2020)
19. He, K.M., Sun, J., Tang, X.O.: Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(12), 2341–2353 (2011)
20. Pérez, P., Gangnet, M., Blake, A.: Poisson image editing. *Acm Trans. Graph.* **22**(3), 313–318 (2003)
21. Jing, Z., Xiao, G., Li, Z.: *Image Fusion: Theory and Application*. Higher Education Press, Beijing (2007)
22. Yates, P.A., Lai, M., Yi, T.W., Martinez, A.: Wide Field Fundus Camera with Auto-montage at a Single Alignment. US Patent: US2019/0159673 A1, PCT/US2017/038560. 30 May 2019
23. Talbar, S.N., Chavan, S.S., Pawar, A.: Non-subsampled complex wavelet transform based medical image fusion. In: Arai, K., Bhatia, R., Kapoor, S. (eds.) *Advances in Intelligent Systems and Computing*, vol. 880, pp. 548–556. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-02686-8_41
24. Fu, J., Li, W., Du, J., et al.: A multiscale residual pyramid attention network for medical image fusion. *Biomed. Signal Process. Control* **66**(9) (2021)