Proceeding of the IEEE
International Conference on Robotics and Biomimetics
Dali, China, December 2019

BAIFA: A Brightness Adaptive Image Fusion Algorithm for Robotic Visual Perception

Jia Cai, Bi Zeng
School of Computer
Guangdong University of Technology
Guangzhou,Guangdong,China
zb9215@gdut.edu.cn

Jianqi Liu, Xiuwen Yin
School of Automation
Guangdong University of Technology
Guangzhou,Guangdong,China
liujianqi@ieee.org

Li He
School of Electromechanical Engineering
Guangdong University of Technology
Guangzhou,Guangdong,China
heli@gdut.edu.cn

***Abstract*— In robot tasks, camera is widely used for scene recognition and localization. However, it is still a challenging problem for robot vision working well in low-brightness environments. We propose a brightness adaptive image fusion algorithm (BAIFA) which fuses one RGB image and one infrared image to improve the quality of image for robotic visual perception. A weight function is presented to calculate the image brightness weight to balance RGB and infrared images in fusion. To verify the proposed algorithm, comparisons with three image fusion algorithms are made in different brightness environments. Experimental results show that the proposed BAIFA is able to effectively preserve image contrast and target contour, which is more robust than others in various brightness environments. Furthermore, a case study shows that, visual perception is improved with our method and the fused image can also provide visual data in mobile robots.**

***Index Terms*— *image fusion, brightness adaptation, image saliency, robotic visual perception*.**

I. Introduction

Visual cameras is widely used for mapping and environmental perception in robotics due to its low cost and rich image information [1]. However, they are poorly adaptable when the scene is poorly lit [2], such as in office, corridor in the evening. Since the image intensities are low, it is difficult to extract sufficient features, recognize targets or separate objects from background [3]. Therefore the stability of the visual system can not be preserved in complex environments.

One way to improve the perception in mobile robot of visible (RGB) image is to integrate an infrad (IR) image [4]. Due to differences in the imaging principles, RGB image and IR image are different even in the same scene. On the one hand, the RGB image clearly presents a high spatial resolution. As a result, it can describe the details of the scene, which is well matched to human visual system [5]. However, it is limited by illumination changes and occlusion. Infrared camera, on the other hand, has strong night recognition capabilities and large target recognition capabilities for complex scenes. However,the IR image lacks detailed information.

In recent years, some researchers have studied the fusion of IR and RGB images [6]–[9], and have made some improvements for vision-based mobile robots [10]. It is shown in many previous works that RGB-IR image fusion technology can obtain more suitable results for human vision, and is convenient for application to detection and recognition. Rong et al. [11] make full use of the complementary information to obtain fused images that are conducive to visual observation or image recognition, localization, and other scenes. The most classic image fusion algorithms are based on wavelet transform, contour transform, or non-subsampled contourlet transform (NSCT) [12]. Visual saliency detection is also used in image processing [13]. It focuses on extracting important target information in images via analysis and then suppresses background and noise [14].

Research on fusion algorithms has rapidly developed in recent years. Most of these results fuse images directly or based on image contrast. However, in vision-based robots ,two specific problems should be solved. 1) The effect of brightness is ignored. The relationship between brightness and image pixel value is not considered in image fusion algorithms, which results in low performance in low brightness case. 2) There is excessive processing of details of fused images, which is time consuming and makes these algorithms unsuitable for robot applications. For a mobile robot, a brightness adaptive image fusion algorithm (BAIFA) is proposed, which considers the influence of brightness change in robot vision processing.

To improving the image fusion quality under complex brightness conditions for robot visual perception, the main contributions in this paper can be summarized as follows:

1)A weight strategy is proposed to balance the importance of RGB and IR components. We balance the weight of the RGB and IR image fusion in different brightness and further split images into base and detail layers. In addition, an adaptive image fusion algorithm is proposed according to the brightness information and salient features of the image. As a result, the overall contrast is preserved and the targets are hence easy to be detected.

2)Evaluations demonstrate the performance of the proposed BAIFA, and the performance is verified from both appearance and performance metrics. Our experiments in different brightness levels show that BAIFA can obtain good fused images to highlight image targets and preserve RGB image detail. In addition, BAIFA outperforms other image fusion algorithms in various brightness environments. Moreover, the practicality and prospects are verified.

The remainder of this paper is organized as follows: Section

978-1-7281-6321-5/19//$31.00 ©2019 IEEE

II introduces related works. In Section III, the brightness adaptive image fusion algorithm is described. Section IV introduces our own datasets and four evaluation metrics. Section V details the experimental design , an analysis of the results,and case studies. Finally, Section VI are Conclusion.

II. Related Works

Image fusion algorithms play a key role in visual perception, and many researchers are studying in this field. The traditional wavelet transform method can effectively extract the multi-scale information of the image. However, it is isotropic, and its sampling operation lacks translation invariance, which leads to inaccuracies in obtaining the edge direction information of the image. NSCT [18], [19] emerged to overcome the above-stated shortcomings, but it only considers the regional characteristics of the high-frequency coefficients. Some problems of the fused images are still difficult to solve in the field of image fusion for robotic visual perception, such as image distortion, lack of detail, and insufficient target.

In a study on the principle of visual saliency, Zhang et al. [20] proposed an infrared feature extraction technique with background removal and reconstruction for low-light environments that can extract important highlight features of IR images for image fusion in a relatively fast time. However, the infrared features are directly superimposed on the RGB image, and the analysis of the difference between the two image features and the processing of the image details are not considered. Li et al. [21] were the first to introduce potential latent low-rank representation (LatLRR) into the field of image fusion. Li et al. decomposed the source image into low-rank and significant parts, and respectively fused them to obtain the final fused image. The fused image is smooth, but the saliency compensation from the IR image is not very prominent, and the algorithm has a long calculation time and cannot be applied well to mobile robot technology.

The principle of visual saliency detection is becoming increasingly mature in image fusion. In the various image fusion methods, it is usually necessary to first decompose the image into a low-frequency sub-image (base layer) and a high-frequency sub-images (detail layer). Based on the rolling Guidance filter (RGF) [22] and the max-absolute rule [23] techniques, Ma et al. used image decomposition techniques based on rolling-guided filtering, which uses the visual saliency map (VSM) for base-layer image fusion and optimizes the principle literature using weighted least square (WLS) [24]. The detail-layer image fusion method preserves the overall appearance and contrast information in the original image and is more in line with the results of human visual observation. For complex brightness scenes, Ma et al.s VSM_WLS algorithm has a good effect on target's highlights, but it still does not fully consider the influence of illumination on image features. Therefore, the proposed method mainly performs brightness adaptive optimization based on the literature [24] and proposes brightness level division and a brightness weight function. The fusion of the base layer has been improved to better adapt to image recognition in various brightness environments, such as darkness, dimness, normal illumination, and brightness.

Overall, this work differs from previous related work in two main aspects: l) we propose an calculation method of brightness weight coefficient. 2) we present an optimized image fusion algorithm to achieve adaptive brightness through base layer fusion and detail layer fusion.

III. The Proposed Algorithm

Generally, it is convenient to use a visual camera for environmental perception in bright environments. However, in poor brightness environments, an additional sensor and fusion algorithm are needed to complement the visual camera and enhance the visual perception. In this paper, a brightness adaptive image fusion algorithm is proposed, with an IR camera, to improve the visual perception performance. The overall framework of BAIFA is shown in Fig. 1.

BAIFA is divided into three portions:

1)The first component preprocesses the images, including grayscale conversion, image enhancement, Gaussian filtering, and image registration. The input RGB image and IR image are separately processed into same-size images that are suitable for the image fusion operation.

2)The second component calculates the brightness parameter of the RGB image. By using the average grayscale image, a brightness function is designed to obtain the brightness weight. Details are described in Section III-A.

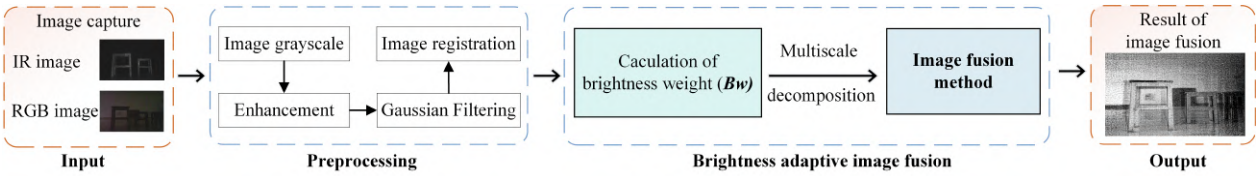
3)The third component is the brightness adaptive image fusion, which includes three processes: image multi-scale decomposition, base layer image fusion, and detail layer image fusion. The base layer fusion preserves the overall contrast and the detail layer fusion retains contour at multiple scales. Finally, the image fusion result is obtained by combining the fused image of the base layer and the plurality of detail layers. The details are described in Section III-B.

A. Brightness Weight Calculation

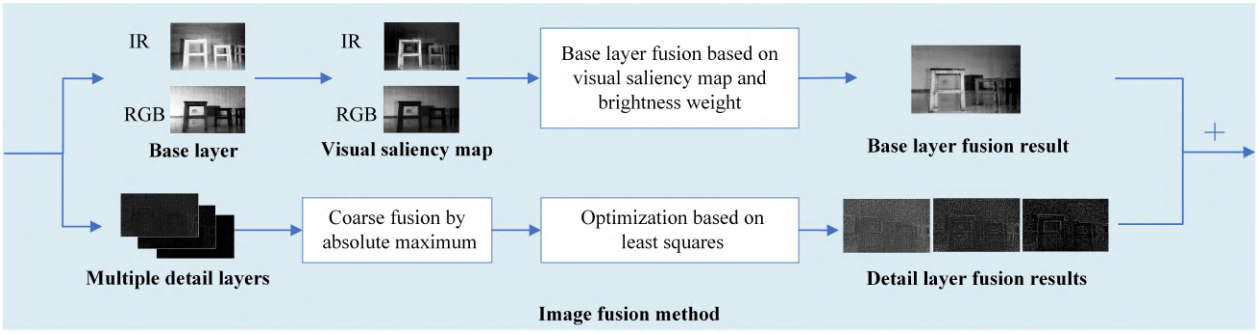
To cope with image fusion under various brightness, the proposed BAIFA algorithm calculates brightness weights and fuses RGB and IR images adaptively. This section focuses on the calculation of image brightness weights.

In this paper, the brightness weight is decided by the RGB image. It can be seen that the overall grayscale value is proportional to the brightness. When the brightness is low, the grayscale value of the image is low. Instead, the grayscale value is high. We discussed the relationship between grayscale values and brightness level for verification, and the result is shown in Fig.2.

In Fig.2, X-axis represents the brightness level. In our verification, 100 indoor RGB images captured from five different times are classified into 5 levels. The Y-axis represents the average grayscale value of the image, we normalize the grayscale value to 0 - 1. More intuitively, one image is chosen from each level as an example. It can be seen from the figure that as the brightness becomes darker, the average grayscale value of the image gradually decreases.



(a)Framework of the proposed BAIFA,including proprocesing and brightness advaptive image fusion way.



(b)The detail image fusion method of BAIFA ,including fusions of base layer and dtail layers.

Fig. 1. The framework of BAIFA.

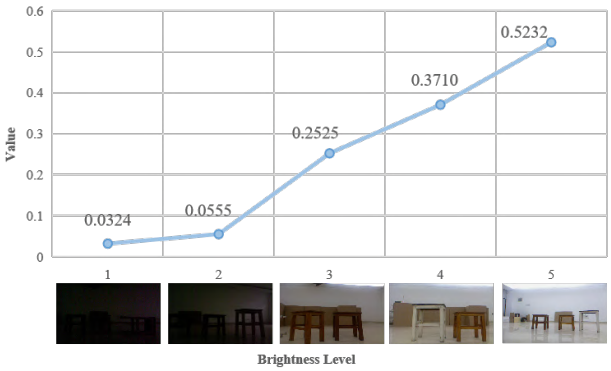


Fig. 2. Relationship between brightness level and normalized grayscale value. The X-axis represents the brightness level and the Y-axis represents the normalized image average brightness value.

In this paper, a method for calculating the brightness parameter is proposed. We define Bw as one of the weight parameters for RGB and IR base layer image fusion. The brightness parameter is defined in (1).

$$Bw = 1 - \log\left(1 + e^{-t}\right), \tag{1}$$

where t represents the average grayscale value of the image pixel.

By deriving t, Bw is increased when the average grayscale value of the RGB image is increased. In addition, (1) causes Bw to fluctuate around 0.5, which can combine the saliency of the infrared image while retaining the RGB information.

Therefore, of one scene, with bright scene, Bw becomes larger, and the fusion weight of the RGB image part increases. When the brightness is low, Bw will decrease.

The pseudo-code for the calculation of the brightness weight is shown in Algorithm 1.

Algorithm 1 Brightness weight of BAIFA

Require: $image\_rgb$
Ensure: Bw
Image ← rgb2gray(*Image_rgb*)
Size ← size(*Image_rgb*)
Count ← hishogram(*Image*)
Sum ← sum(*Count*) //calculate total number of all pixels
//Calculation of brightness weight parameters Bw
for i ← 0 to 1: **do**
    //Product of the i pixel and its pixel amount
    $Tmp \leftarrow i * Count(i)$
end for
$Sum\_tmp \leftarrow sum(Tmp)$
$t \leftarrow Sum\_tmp/(255 * Sum)$
$Bw \leftarrow 1 - \log(1 + e^{-t})$

B. Brightness Adaptive Image Fusion

1) Multi-Scale Decomposition: The rolling guidance filter (RGF) algorithm has both scale-aware and edge-preserving properties [22]. It separates the large-scale edge, small-scale texture, and the underlying coarse-scale information of the image into a base layer and multiple detail layers. RGF includes two main steps: small structure removal and edge recovery. Firstly, in order to achieve denoising and de-interference, Gaussian filtering is employed to remove small areas of the image, and a process eliminating structures with small scales from the image. Secondly, iterative edge recovery ensures the accuracy of the large area boundary. The flowchart of image decomposition based on RGF is shown in Fig. 3.

Then, the input image can be decomposed into the base layer B and the detail layers $d^1, d^2, ..., d^N$. The multi-scale decomposition (MSD) based on RGF may be suitable for the

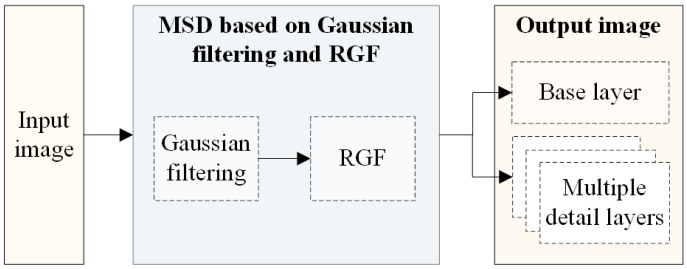


Fig. 3. The flow of image decomposition based on RGF.

application of a multi-scale image fusion, since they can well achieve scale separation of spatially-overlapped features. The base layer mainly retains the global contrast and appearance of the image. The detail layer retains details and textures, which can reflect the contour details of the image and remove the complicated background.Meanwhile, it overcomes the halo problem near the edges.

2) Base Layer Fusion Based on Brightness Weight and VSM: In this paper, the principle of visual saliency map (VSM) has been used to perform the grassroots fusion [13]. The fusion weight of the base layer can be obtained according to the contrast value of each pixel, and the brightness of the current environment. The base layer of the RGB image and the infrared image is adaptively fused through saliency detection.

For the basic layer of the RGB and IR images, the contrast of each pixel value is calculated. Then, pixel with the same intensity are accumulated to obtain a saliency for each pixel value. Defining the saliency value of pixel i is $v(i)$, then:

$$v(i) = \sum_{j=0}^{255} P_j |I_i - I_j|, \tag{2}$$

where P_j represents the number of pixels whose pixel values are equal to j. Finally, the saliency maps v_1 and v_2 RGB image and the IR image respectively, are obtained.

Through the image saliency map and brightness weight, the base-layer fusion of the RGB image and the IR image is carried out, and the base-layer fusion image BF is defined as follows:

$$\begin{aligned} BF =& Bw \cdot [1 + (v_1 - v_2)] \cdot B_1 \\ &+ (1 - Bw) \cdot [1 + (v_1 - v_2)] \cdot B_2, \end{aligned} \tag{3}$$

where Bw denotes the brightness weight; B_1 and B_2 represent the primary images of RGB image and infrared image respectively; and v_1 and v_2 represent the saliency map results respectively.

The fusion rules of the base layer are provided for the base layer fusion by considering the brightness parameter Bw and the results of VSM. In terms of image brightness, the fusion weight of B_1 increases if Bw is larger.

3) Detail Layer Fusion Based on WLS: The detail layer removes complex background interference and retains the contour of the image. An effective fusion of the detail layer can highlight the foreground object of the image, which is beneficial to target recognition.

The traditional absolute maximum rule can fuse the detail layer, i.e. compare the detail layer image of each level of the two images. For the j-level detail layer, d_1 and d_2 respectively represent the pixel values of the detail layer of the RGB and IR images respectively. In this paper, if there exists a strong contour in the infrared image, the fusion coefficient Dw is 1, and 0 otherwise. Correspondingly, the fusion coefficient of RGB image is $(1 - Dw)$.

Infrared image is presented to display a soarse contour. However, it lacks detail and has severe speckle noise, which cause some defects to the fused images. In our paper, after the rough fusion, a loss function is set according to WLS, and the optimal detail layer fusion image D of each level is obtained. According to the loss function from [24], $Loss(D)$ is finally defined as follows:

$$Loss(D) = (D - R)^2 + \lambda | \sum_{q \in win} (d_2^j)_q |^{-1} (D - d_2)^2. \tag{4}$$

The purpose of the first squared term is to make the resulting fused image D^j close to the result of rough fusion while the second is to preserve RGB details as much as possible. λ denotes the balance parameter.In the inverse of an absolute value, pixel sum of the j-th infrared detail layer is obtained, where win is a 7x7 sliding window.

For the j-level detail layer, the extreme value of $Loss(D_j)$ is taken to obtain the minimum D_j.

Finally, F, which represents image fusion results based on brightness adaptation, is composed of the superposition and composition of base-layer fusion BF and multiple detail layers fusion D_j, as shown in (5).

$$F = BF + D_1 + D_2 + \cdots + D_N. \tag{5}$$

IV. Datasets and Evaluation Metric

A. The Datasets

For this study, datasets consists of 400 indoor images in order to evaluate the performance of the proposed method in terms of different brightness conditions. As shown in Fig.4. the first row represents the RGB and the second of IR images. Both were generated by a Kinect-V1 camera. To create different brightness environments, we collected images from five different times (specifically at 16:00, 17:00, 18:00, 19:00, and 20:00, respectively).

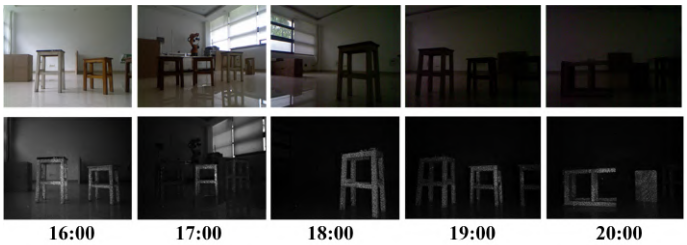


Fig. 4. Different brightness images.

B. Evaluation Metric

To evaluate the performance of image fusion algorithm on the selected datasets, four popular metrics were chosen, including Entropy (E) [25], Root Mean Square Error (RMSE) [26], Edge Protection Index (EPI) [27] and Peak Signal to Noise Ratio (PSNR) [28].

Entropy (E) is an indicator of the richness of image information, which reflects the performance of the image detail and the amount of information carried by the image. A larger entropy value means the image contains more detail information, which is equivalent to better image-fusion quality. RMSE is the standard for evaluating algorithm accuracy. It can represent the overall difference between the two images in image fusion techniques. A lower RMSE indicates that the difference is lower and the algorithm has higher precision.An EPI processing after-filter maintains the horizontal or vertical edge of the image. A higher EPI indicates stronger edge retention of the algorithm, and it is often used to measure the ability for maintaining image detail. PSNR is a similarity evaluation index of the reference image and the result image, which can also reflects the noise condition in the result image. A larger value represents a better fusion effect.

V. Experiments and Analysis

The code runs on a personal computer equipped with Intel(R) Core (TM) i7-6700 model CPU, and 8GB memory, Windows 7 (64-bit) operating system, and MATLAB 2016 simulation software.

A. Comparative Experiments

We compared the proposed BAIFA method with three fusion methods, including the LatLRR [21], the VI_IF [21], and the VSM_WLS algorithm [24]. The results are shown in Fig.5, where the first column is the RGB image, the second column is the IR image, column 3 to 5 are results of competitive, and the last column is the proposed method. From top to bottom, five rows represent five different brightness.

Compared with these three fusion methods, the method achieved the best result between 18:00 and 20:00. BAIFA can highlight the targets and reduce background exposure. Furthermore, it retains more texture information from the entire image. From 16:00 to 17:00, the performances of the VSM-WLS method and BAIFA were similar. It is because when the brightness value is large, Bw of BAIFA in (1) is close to 0.5, which is the same as the weight coefficient of VSM-WLS.

Four metrics (Entropy, RMSE, EPI, and PSNR) were used to evaluate the LatLRR, VI_IF, VSM_WLS and BAIFA methods on the datasets, as shown in Table I. When the brightness rangeds from 20:00 to 16:00, the proposed method achieved the best results. When the brightness is of 16:00, the RMSE and PSNR of VSM_WLS are slightly better than the proposed method. Moreover, a comparison column of the algorithm performance evaluation results is shown in Fig.6.

For the RGB and IR image fusion experiments with a pixel size of 1373 x 1026, an operation time comparison of various algorithms was performed. The average operation time comparison results are shown in Table II.

TABLE I
ALGORITHM PERFORMANCE EVALUATION RESULT.

Image	Metric	LatLRR	VI_IF	VSM_WIS	BAIFA
(a)20:00	Entropy	0.974	0.595	0.995	**0.996**
	RMSE	0.245	0.325	0.250	**0.223**
	EPI	0.731	1.111	1.909	**1.928**
	PSNR	60.34	57.88	60.14	**61.23**
(b)19:00	Entropy	0.939	0.527	0.989	**0.994**
	RMSE	0.261	0.353	0.238	**0.222**
	EPI	0.739	1.200	2.319	**2.326**
	PSNR	59.78	57.16	60.59	**61.17**
(c)18:00	Entropy	0.921	0.513	0.988	**0.997**
	RMSE	0.241	0.3820	0.268	**0.246**
	EPI	0.866	1.054	1.953	**1.946**
	PSNR	60.47	56.48	59.68	**60.59**
(d) 17:00	Entropy	0.987	0.468	0.997	**0.999**
	RMSE	0.259	0.362	0.255	**0.253**
	EPI	1.282	1.704	3.905	**3.921**
	PSNR	59.83	56.98	59.98	**60.05**
(e)16:00	Entropy	0.986	0.671	0.990	**0.993**
	RMSE	0.248	0.346	**0.199**	0.201
	EPI	1.110	1.592	3.239	**3.284**
	PSNR	60.28	57.33	**62.18**	62.13

TABLE II
A COMPARISON OF THE AVERAGE OPERATION TIME.

Algorithm	LatLRR	VI_IF	VSM_WIS	BAIFA
Average time/s	1037.7360	2.0663	4.2700	4.2704

According to the operation time of each algorithm in Table II, BAIFA , VSM_WLS and VI_IF, these three methods show a similar time cost in our test and is much faster than LatLRR.

In general, the proposed algorithm combines saliency features and contour details of the IR image on the target object, and thus it can highlight the target and reduce misidentification. In a complex brightness background, especially in the case of insufficient brightness, when the single visual camera cannot work, there are obvious effects and the advantage is more prominent. Therefore, the proposed BAIFA can provide a good image for object recognition, and is useful for robot visual perception in different brightness environments.

B. Case Study

To better verify the performance of the algorithm, experiments were performed on public datasets. We applied our proposed algorithm in vehicle and human identification in complex outdoor environments. Two cases of the fusion results from public dataset images were used, as showed in Fig.7.

It can be seen that the algorithm can also be used for outdoor identification. RGB images cannot be used to identify

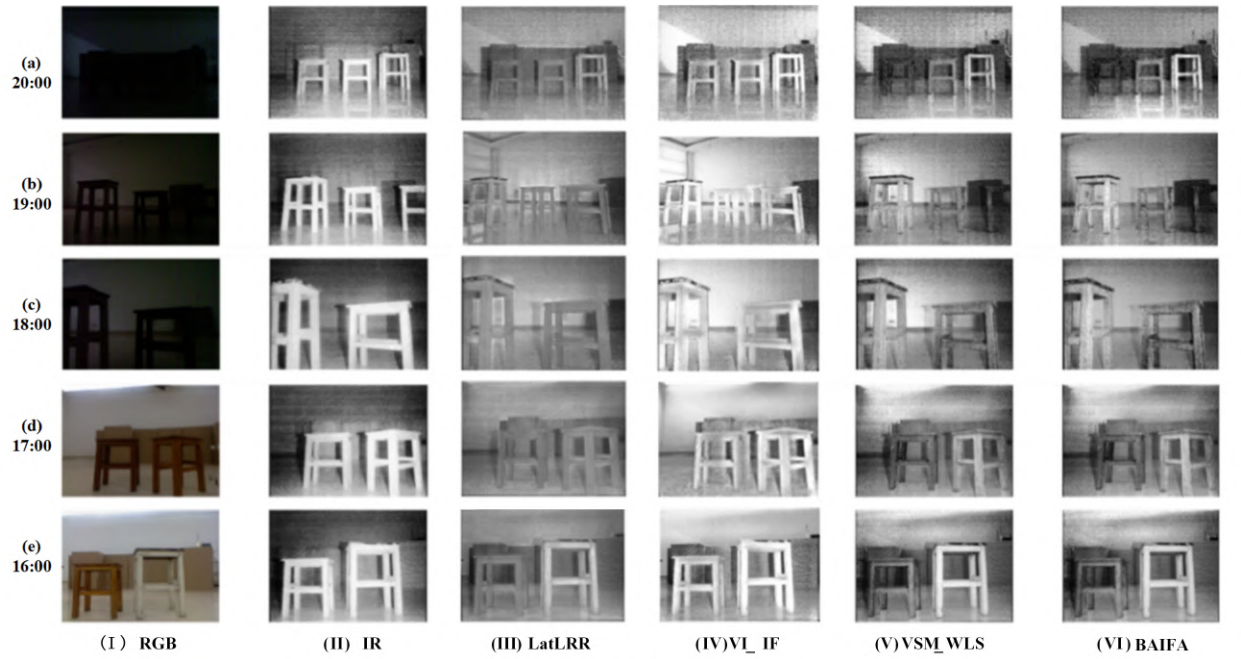


Fig. 5. A comparison of the fusion results from different methods. In each group, we present (I)the RGB images acquired by Kinect camera, (II)the preprocessed IR image, (III) the fusion result images of the LatLRR algorithm, (IV) the fusion result images of the VI_IF algorithm , (V) fusion result images of original VSM_WLS algorithm , and (VI) fusion result images of BAIFA.

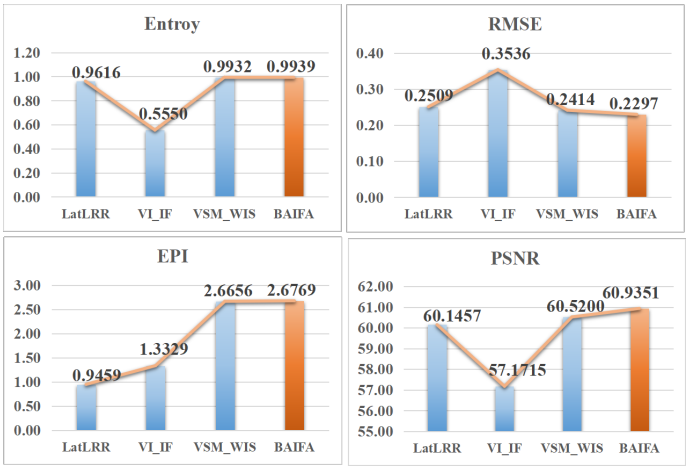


Fig. 6. A comparison column of algorithm quality evaluation results. In the diagram, the orange bar represents the proposed BAIFA. The higher the value of Entroy, EPI and PSNR represent the better image fusion. Conversely, the smaller the value of RMSE is the better.

Fig. 7. The fusion results of BAIFA from two public dataset images.

TABLE III
BAIFA CALCULATION TIME STATISTICS OF DATA SET IMAGES.

Image	Marne	Kaptein
Time/s	2.8132	2.7121

vehicles and humans before image fusion due to the complex brightness, similar colors, and other problems, all of which have a great impact on safety, and the potential risks reduce the application of patrol robots. With the proposed algorithm, a fused image can be used for robot visual recognition and where the fused image relatively clear. Furthermore, the calculation time of the dataset image fusion algorithm is shown in Table III.

Since the data set image has been pre-processed through compression and other processes, the image resolution is low and the fusion time is also low, which can basically meet the requirement of real-time processing.

The case study verifies the effectiveness of our algorithm on public datasets. After using the proposed method, the vehicle and passersby can be clearly restored and highlighted. This shows the practicality and security of robotic visual perception.

VI. Conclusion

In this paper, a brightness adaptive image fusion algorithm (BAIFA) is proposed to improve the adaptability under differ-

ent brightness environments. The fusion threshold of our proposed BAIFA is changed with the different brightness weight. The experimental results show that our method performs better than LatLRR, VI IF and VSM WLS, three state-of-the-art methods. In addition, the BAIFA not only preserves the texture details of the RGB image, but also can emphasize the targets.

In the future, our research will focus on the computation time reducing . In addition, the application of merged image in scene recognition and robot localization are directions for future research works.

ACKNOWLEDGMENT

This work was supported in part by grants from the National Natural Science Foundations of China (Nos. 61701122, 61703115), the Natural Science Foundation of Guangdong Province, China (Nos. 2016A030313734, 2018A030310540, 2017A030313431), Science and Technology Program of Guangzhou, China (No. 201804010238).

A. References

REFERENCES

[1] D. M. Tsai, W. Y. Chiu, and T. H. Tseng, "Moving object detection from a mobile robot using basis image matching," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 9406, no. 1, pp. 1–10, 2015.

[2] Q. K. Dang and Y. S. Suh, "Human-following robot using infrared camera," in *2011 11th International Conference on Control, Automation and Systems*. IEEE, 2011, pp. 1054–1058.

[3] Y. Zuo, J. Liu, G. Bai, X. Wang, and M. Sun, "Airborne infrared and visible image fusion combined with region segmentation," *Sensors*, vol. 17, no. 5, p. 1127, 2017.

[4] H. Li, B. Wang, and L. Li, "Research on the infrared and visible power-equipment image fusion for inspection robots," in *2010 1st International Conference on Applied Robotics for the Power Industry*. IEEE, 2010, pp. 1–5.

[5] J. Ma, Y. Ma, and C. Li, "Infrared and visible image fusion methods and applications: A survey," *Information Fusion*, vol. 45, pp. 153–178, 2019.

[6] H. Ghassemian, "A review of remote sensing image fusion methods," *Information Fusion*, vol. 32, pp. 75–89, 2016.

[7] Y. Kurmi and V. Chaurasia, "An image fusion approach based on adaptive fuzzy logic model with local level processing," *International Journal of Computer Applications*, vol. 124, no. 1, 2015.

[8] Y. Chen, J. Xiong, H.-l. Liu, and Q. Fan, "Fusion method of infrared and visible images based on neighborhood characteristic and regionalization in nsct domain," *Optik-International Journal for Light and Electron Optics*, vol. 125, no. 17, pp. 4980–4984, 2014.

[9] X. Jin, Q. Jiang, S. Yao, D. Zhou, R. Nie, J. Hai, and K. He, "A survey of infrared and visual image fusion methods," *Infrared Physics & Technology*, vol. 85, pp. 478–501, 2017.

[10] A. Toet, M. A. Hogervorst, S. G. Nikolov, J. J. Lewis, T. D. Dixon, D. R. Bull, and C. N. Canagarajah, "Towards cognitive image fusion," *Information Fusion*, vol. 11, no. 2, pp. 95–113, 2010.

[11] C. Rong, Y. Jia, Y. Yang, Y. Zhu, and Y. Wang, "Fusion of infrared and visible images through a hybrid image decomposition and sparse representation," in *2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, vol. 2. IEEE, 2018, pp. 21–25.

[12] S. Das and M. K. Kundu, "Nsct-based multimodal medical image fusion using pulse-coupled neural network and modified spatial frequency," *Medical & biological engineering & computing*, vol. 50, no. 10, pp. 1105–1114, 2012.

[13] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*. Ieee, 2007, pp. 1–8.

[14] T. N. Vikram, M. Tscherepanow, and B. Wrede, "A saliency map based on sampling an image into random rectangular regions of interest," *Pattern Recognition*, vol. 45, no. 9, pp. 3114–3124, 2012.

[15] J. Han, E. J. Pauwels, and P. De Zeeuw, "Fast saliency-aware multi-modality image fusion," *Neurocomputing*, vol. 111, pp. 70–80, 2013.

[16] H.-x. Liu, T.-h. Zhu, and J.-j. Zhao, "Infrared and visible image fusion based on region of interest detection and nonsubsampled contourlet transform," *Journal of Shanghai Jiaotong University (Science)*, vol. 18, no. 5, pp. 526–534, 2013.

[17] J. Adu, S. Xie, and J. Gan, "Image fusion based on visual salient features and the cross-contrast," *Journal of Visual Communication and Image Representation*, vol. 40, pp. 218–224, 2016.

[18] Y. H. Wu, D. Yan, M. X. Ma, and N. Wu, "An improved compressive sensing image fusion algorithm based on nsct transform," in *Applied Mechanics and Materials*, vol. 433. Trans Tech Publ, 2013, pp. 306–309.

[19] Y. Yang, Y. Que, S. Huang, and P. Lin, "Multimodal sensor medical image fusion based on type-2 fuzzy logic in nsct domain," *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3735–3745, 2016.

[20] Y. Zhang, L. Zhang, X. Bai, and L. Zhang, "Infrared and visual image fusion through infrared feature extraction and visual information preservation," *Infrared Physics & Technology*, vol. 83, pp. 227–237, 2017.

[21] H. Li and X.-J. Wu, "Infrared and visible image fusion using latent low-rank representation," *arXiv preprint arXiv:1804.08992*, 2018.

[22] Q. Zhang, X. Shen, L. Xu, and J. Jia, "Rolling guidance filter," in *European conference on computer vision*. Springer, 2014, pp. 815–830.

[23] O. Prakash, A. Kumar, and A. Khare, "Pixel-level image fusion scheme based on steerable pyramid wavelet transform using absolute maximum selection fusion rule," in *2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*. IEEE, 2014, pp. 765–770.

[24] J. Ma, Z. Zhou, B. Wang, and H. Zong, "Infrared and visible image fusion based on visual saliency map and weighted least square optimization," *Infrared Physics & Technology*, vol. 82, pp. 8–17, 2017.

[25] X. Luo, J. Zhang, J. Yang, and Q. Dai, "Image fusion in compressed sensing," in *2009 16th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2009, pp. 2205–2208.

[26] V. Vijayaraj, C. G. O'Hara, and N. H. Younan, "Quality analysis of pansharpened images," in *IEEE International Geoscience & Remote Sensing Symposium*, 2004.

[27] S. Gupta, L. Kaur, R. C. Chauhan, and S. C. Saxena, "A versatile technique for visual enhancement of medical ultrasound images," *Digital Signal Processing*, vol. 17, no. 3, pp. 542–560, 2007.

[28] D. M. Chandler and S. S. Hemami, *VSNR: a wavelet-based visual signal-to-noise ratio for natural images*, 2007.