2020 19th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)

Visible and infrared image fusion based on visual saliency detection

Xizi Tan
School of Computer Science and Technology
Huaiyin Normal University
Huai'an,China
1070912428@qq.com

Liqiang Guo
School of Computer Science and Technology
Huaiyin Normal University
Huai'an,China
guolq@hytc.edu.cn

***Abstract*—Visible and infrared image fusion is widely adopted in navigation, object recognition and smart city. In this paper, we use the saliency object detection method to fusion the visible and infrared images. In consideration of the properties of the different type of image, we design two different visual saliency object detection methods to fuse the coefficients of detail image layers and coarse image layers. Then we can reconstruct the fused image from those coefficients. The experimental results indicate that the proposed fusion algorithm conforms to the human vision characteristics and has a strong comparability and effectiveness in both objective and subjective evaluation.**

***Keywords*-Image fusion; Saliency object detection; Morphological transform.**

I. Introduction

Image fusion has been developed for nearly 50 years, which can synthesize the useful information from different sensors into the fused image. As an important image processing method, image fusion is widely used in the domain of object detection, computer vision, smart city and navigation. The virtue of image fusion in that it can provides more information than individual captured image. Among the category of image fusion, visible and infrared image fusion is one of the most important branches. For the low contrast imaging conditions, the image captured by the traditional visible sensor can not provide more sufficient information about the sense. In contrast, the infrared image is captured by the heat emission based on the thermal radiation of the target. As a result, the infrared image can capture the hidden target. For the purpose of detection of a target, the image fusion technique can provide a better solution.

Recently, the visible and infrared image fusion algorithm has made great progress. Considering the complementary prosperity of the visible and infrared image, the existing fusion methods mainly belongs to spatial domain and frequency domain (i.e. transformation based method). The spatial domain methods include: average or weight algorithm, principle component analysis (PCA), data segmentation algorithm, focus measure algorithm [22] and singular value decomposition (SVD). As for the detailed implementation steps about these fusion methods, we recommend the author to consult Mitchells book [1].

For the frequency domain methods, we use specific transformation to convert the source image from spatial domain to the frequency domain, and then use the fusion rule to obtain the final coefficients. Last, we can use the inverse transformation on those coefficients to get the fused image [2]. The discrete wavelet based algorithm [3] is one of the most important one belongs to this category. Some hybrid methods based on the wavelet transform is also utilized, such as incorporating the HSI or HSV transform into wavelet methods, the combination of PCA and wavelet [4]. Besides, the Laplacian pyramid [5], discrete curvelet transform [6], the NSCT [7] and the NSST [8] based methods also belong to this category.

In recent years, some excellent fusion algorithm was proposed. Liu et.al. proposed the adaptive sparse representation (ASR) based fusion algorithm[9]. The combination of anisotropic diffusion and Karhunen-Loeve transform for the purpose of fusion [10]. Besides, the fourth order partial differential equation is also adopted in the image fusion [11]. The author in reference [12] use the multi-level Gaussian curvature filtering (MGCF) to construct an effective fusion algorithm. Anantrasirichai et.al. proposed the fusion algorithm based on Sparse Regularization with Non-Convex Penalties [13]. Besides the paper mentioned above, there are also a lot of excellent algorithm was proposed. Due to the space limit, we recommend the reader to read the review article [14-15].

The goal of multi-sensor image fusion algorithm is to preserve the identifiability of infrared object in the scene observed by human. Ascribe to the sensibility of human eyes to overall contrast information, the significance of image fusion is infrared target and contrast information. To prevent the loss of contrast and detail information, we design two different visual saliency detection algorithms for high and low frequency image layers. The contribution of this paper is that we proposed the simple and effective method for infrared saliency feature extraction, which is crucial for multi-sensor image fusion.

II. Proposed method

For the image fusion, one of the most important things is to detect the useful information and enhance the sharpness information. Based on this consideration, we use the visual saliency object detection, which have been constantly studied in different domains, such as remote sensing, biomedical imaging and video monitoring [16]. The human visual system can capture and locate the saliency object in a natural scene. The visual saliency detection algorithm, by simulating the human visual characteristics, is a method to

2473-3636/20/$31.00 ©2020 IEEE
DOI 10.1109/DCABES50732.2020.00043

find the significant objects in the infrared image. Based on those considerations, we give two visual saliency detection algorithms for the base (or coarse) and detail image layer, and proposed the visual saliency object detection based visible and infrared image fusion algorithm.

A. Fusion framework

The proposed algorithm flow chart is shown in Fig.1. In the first step, we select the Gaussian filter to decompose the visible and infrared images, and get the coarse and detail image layer [17]. In consideration of the noise in the sensor images (especially for the infrared image), we select a linear smoothing filter (Gaussian filter) to suppress the noise.

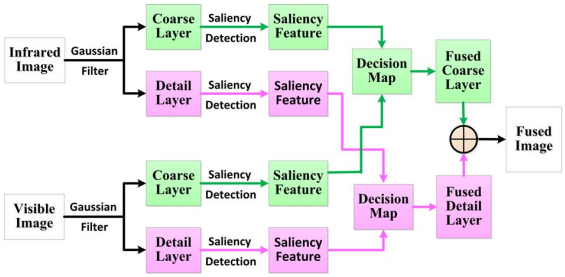


Figure 1: flow char of the proposed fusion method

We use the Gaussian filter to decompose the source image into base (or coarse) image layer and detail image layer, this can be represented by the following equation

$$C_{IR} = I_{IR} * G \tag{1}$$

$$C_{VI} = I_{VI} * G \tag{2}$$

where C_{VI} and C_{IR} are the coarse image layer of the visible image I_{VI} and the infrared image I_{IR}, respectively; the symbol "G" represent the rotationally symmetric Gaussian low-pass filter of size 11 with standard deviation 5. The symbol "$*$" represent the convolution operation.

The detail image layers are extracted by the following equation

$$D_{IR} = I_{IR} - C_{IR} \tag{3}$$

$$D_{VI} = I_{VI} - C_{VI} \tag{4}$$

where D_{VI} and D_{IR} are the detail image layers of the visible and infrared image, respectively.

In the second step, the saliency object detection algorithm was adopted to find the significant region. We know that the detail layer contains high-frequency information such as the region boundary or edge of the original image, while the coarse layer contains the approximate structure and contrast information. In view of their different features, we adopt two different visual saliency detection methods for the coarse and detail layers.

In the third step, based on the saliency feature of coarse and detail image layer, we use the weighted averaging method to get the decision map. Then we can obtain the fused coarse and detail image layer.

In the last step, we use the linear combination of coarse image layer and detail image layer to obtain the final fused image.

B. Fusion of coarse layer

For the coarse layer fusion, it is important to extract edge information, improve the contrast of images and highlight the infrared object. Here, we present a saliency detection algorithm which can enhance the contrast information. The method determines the salient value by comparing each pixel with the mean of coarse layer. Take the infrared coarse layer C_{IR} as an example, the corresponding saliency value is defined as

$$V_{IR} = | I_P - M_{IR} | \tag{5}$$

where M_{IR} denotes the mean value of infrared image I_{IR}, I_P is the intensity value of a pixel in I_{IR}.

Then the saliency map is generated by normalizing the V_{IR}. Compared with the method proposed in [18], this method is more concise and faster, neither requiring the pixels of an image to be within a certain range, nor requiring the intensity value of a point pixel to be compared with all other pixels.

In the same way, the saliency value V_{VI} of visible coarse layer is calculated:

$$V_{VI} = | I_P - M_{VI} | \tag{6}$$

where M_{VI} represents the mean value of visible image I_{VI}.

We define the decision map as:

$$w_1 = 0.5 + (V_{IR} - V_{VI}) \times 0.5 \tag{7}$$

$$w_2 = 1 - w_1 \tag{8}$$

Then, the final fused coarse layer is:

$$C_f = w_1 \cdot C_{IR} + w_2 \cdot C_{VI} \tag{9}$$

where the symbol "$\cdot$" represent the point multiplication.

C. Fusion of detail layer

In the process the fusion detail image layers, we recommends to remove the image noise while preserving edges and detail information as much as possible, so as to get a complete and clear visual saliency map. Here, we select the morphological algorithm to generate visual saliency map of detail layer.

For visible image, we first carry out the morphological top-hat transform on the original detail layer image D_{VI}. In other words, we calculate the morphological opening operation and then subtract the result from the original image. For the detailed implementations of the top-hat transform, we select a disk-shaped structuring element, where the radius is 15.

Next, we perform the morphological bottom-hat transform on the original detail layer image D_{VI}. Here, we use similar disk-shaped structuring element mentioned above for the implementation of bottom-hat transform.

Finally, the saliency feature of the original detail layer image D_{VI} is obtained by the subtraction of the result of top-hat transform and the bottom-hat transform.

In the same way, we can use the above procedure to obtain the saliency feature of the original infrared detail layer D_{IR}. It should be noted that the structuring element is no longer the disk-shaped one, but a square structuring element whose width is 15 pixels.

Then, the fused detail layer D_f is obtained by the maximum selection rule.

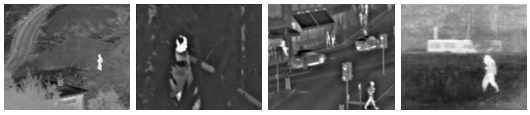
D. Image reconstruction

Finally, combine the fused detail layer D_f and coarse layer C_f, the fusion image F is obtained by the following equation:

$$F = C_f + D_f \tag{10}$$

III. Experimental results and discussion

In this section, we carry out an experiment on the four pairs of source image from the dataset [19], as shown in Fig .2 and Fig .3. Besides, the traditional fusion algorithm and recent developed methods are selected to compare with the proposed method. Those fusion methods including discrete wavelet transform (DWT), gradient pyramid(GP) [20], the adaptive sparse representation (ASR) based method [9], the anisotropic diffusion and Karhunen-Loeve transform (ADKL) based method [10], the partial differential equations (PDE) based method [11], the multi-level Gaussian curvature filtering (MGCF) based method [12] and the Sparse Regularization with Non-Convex Penalties (SRNP) based method [13]. The experimental results are listed in Fig. 4 - Fig. 7.



(a) Camp (b) Tree (c) Road (d) Forest

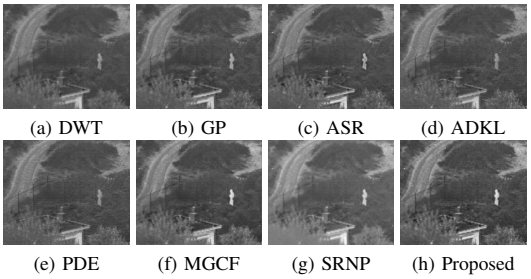
Figure 2: The infrared image.

(a) Camp (b) Tree (c) Road (d) Forest

Figure 3: The visible image.

A. Subjective evaluation analysis

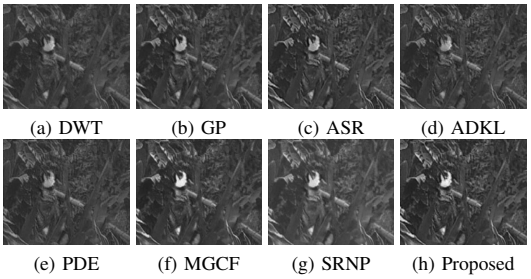
From Fig. 4 - Fig. 7 , we can find that the DWT, the ADKL and the PDE based method exhibit low contrast, and the detailed information is relative poor. The ASR and the SRNP based method preforms second best, and those two methods behave relative better visual perception. The MGCF and the proposed method perform best among those methods, the salient infrared target and detailed information is more superior to other methods.



(a) DWT (b) GP (c) ASR (d) ADKL

(e) PDE (f) MGCF (g) SRNP (h) Proposed

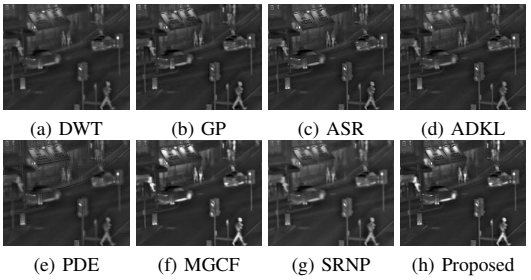
Figure 4: The fusion result of the "Camp" image.



(a) DWT (b) GP (c) ASR (d) ADKL

(e) PDE (f) MGCF (g) SRNP (h) Proposed

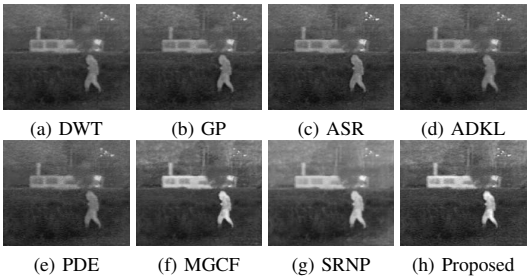
Figure 5: The fusion result of the "Tree" image.



(a) DWT (b) GP (c) ASR (d) ADKL

(e) PDE (f) MGCF (g) SRNP (h) Proposed

Figure 6: The fusion result of the "Road" image.



(a) DWT (b) GP (c) ASR (d) ADKL

(e) PDE (f) MGCF (g) SRNP (h) Proposed

Figure 7: The fusion result of the "Forest" image.

B. Objective evaluation analysis

The objective evaluation metrics are: average gradient (AG), spatial frequency (SF), information entropy (EN), and visual information fidelity index (VIF) [21]. The corresponding objective evaluation results are listed in Table I, where the results in bold face indicate better fusion results. The experimental data in Table I indicate that the fusion images of the proposed algorithm can retain more detail information, which is obviously better than the typical and recent developed methods.

Table I: Objective assessment of different fusion methods

Images	Metric	DWT	GP	ASR	ADKL	PDE	MGCF	SRNP	Proposed
Camp	AG	2.7554	3.5500	3.5011	3.3602	3.4679	4.1529	3.1103	**4.7390**
	Entropy	6.2441	6.3500	6.3362	6.2865	6.2887	6.5973	6.5309	**6.6401**
	SF	6.5898	9.6126	9.3359	8.0915	8.1801	11.3646	7.7000	**12.7058**
	VIF	0.3228	0.3308	0.3178	0.2794	0.2835	0.4359	0.3793	**0.4560**
Tree	AG	2.6248	3.4449	4.0471	3.6407	3.3141	3.9005	3.3985	**4.5574**
	Entropy	5.9426	6.1740	6.2281	6.0887	6.0201	6.2960	6.2829	**6.4010**
	SF	5.4209	7.3373	8.7083	7.6530	6.8155	9.0490	7.3769	**10.2780**
	VIF	0.3170	0.3762	0.4224	0.3550	0.3342	0.4707	0.4556	**0.4986**
Road	AG	1.9948	2.6515	2.8405	2.6825	**4.1046**	3.2804	2.2392	3.4891
	Entropy	5.9178	6.0608	6.0456	6.0397	6.0671	6.1198	**6.2728**	6.1351
	SF	6.8573	10.9381	11.8599	9.7897	12.4838	12.7548	7.7711	**13.2100**
	VIF	0.3621	0.4137	0.4278	0.3571	0.2720	0.5545	0.4164	**0.5481**
Forest	AG	1.7512	2.0842	2.1503	2.1869	2.2558	2.6267	1.7960	**2.9676**
	Entropy	6.7824	6.8029	6.8173	6.8057	6.8083	7.0864	**7.1626**	7.1484
	SF	3.5072	4.4012	4.6226	4.4511	4.5276	6.1219	3.8531	**6.6309**
	VIF	0.3967	0.4079	0.4330	0.3773	0.3856	0.5317	0.4070	**0.6043**

IV. Conclusion

This work presents a visible and infrared image fusion algorithm based on visual saliency object detection. The main virtue of the proposed algorithm can be summarized in the following three aspects. First, it can highlight the infrared target and retain relatively complete contrast information. Second, it can effectively reflect the edge details and other information. Third, it exhibit good visual effect. The experimental results show that the proposed algorithm is suitable to the image fusion tasks.

References

[1] H.B.Mitchell, "Image fusion: theories, techniques and applications," Springer, 2010.

[2] L.Guo, X.Cao, L.Liu, "Dual-tree biquaternion wavelet transform and its application to color image fusion," Signal Process., 2020, 171: 107513.

[3] G.Pajares, J.M.Cruz, "A wavelet-based image fusion tutorial," Pattern Recogn., 2004, 37: 1855-1872.

[4] K.Amolins, Y.Zhang, P.Dare, "Wavelet based image fusion techniques - An introduction, review and comparison," ISPRS J. Photogram., 2007, 62: 249-263.

[5] H.Xu, Y.Wang, Y.Wu, Y.Qian, "Infrared and multi-type images fusion algorithm based on contrast pyramid transform," Infrared Phys. Techn., 2016, 78: 133-146.

[6] Z. Shao, J. Liu, Q. Cheng, "Fusion of infrared and visible images based on focus measure operators in the curvelet domain," Appl. Optics, 2012, 51 (12): 1910-1921.

[7] C. Zhao, Y. Guo, Y. Wang, "A fast fusion scheme for infrared and visible light images in NSCT domain," Infrared Phys. Techn., 2015, 72: 266-275.

[8] W. Kong, Y. Lei, H. Zhao, "Adaptive fusion method of visible light and infrared images based on non-subsampled shearlet transform and fast non-negative matrix factorization," Infrared Phys. Techn., 2014, 67: 161-172.

[9] Y. Liu and Z. Wang, "Simultaneous image fusion and denoising with adaptive sparse representation," IET Image Process., 2015, 9(5): 347-357.

[10] D.P.Bavirisetti, R.Dhuli, "Fusion of infrared and visible sensor images based on anisotropic diffusion and Karhunen-Loeve transform," IEEE Sens. J., 2016, 16(1): 203-209.

[11] Z.Zhou, B.Wang, S.Li, et al, "Multi-sensor image fusion based on fourth order partial differential equations," in Int. Conf. Information Fusion, Xi'an, 2017, pp. 1-9.

[12] W.Tan, H.Zhou, J. Song, et al, "Infrared and visible image perceptive fusion through multi-level Gaussian curvature filtering image decomposition," Appl. Optics, 2019, 58(12): 3064-3073.

[13] N.Anantrasirichai, R.Zheng, I.Selesnick, A.Achim, "Image Fusion via Sparse Regularization with Non-Convex Penalties," Pattern Recogn. Lett., 2020, 131: 355-360.

[14] J.Ma, Y.Ma, C.Li, "Infrared and visible image fusion methods and applications: A survey," Inform. Fusion, 2019, 45: 153-178.

[15] X.Jin, Q.Jiang, S.Yao, et.al., "A survey of infrared and visual image fusion methods," Infrared Phys. Techn., 2017, 85: 478-501.

[16] L.Zhang,Q.Sun, "Saliency detection and region of interest extraction based on multi-image common saliency analysis in satellite images," Neurocomputing, 2018, 283:150-165.

[17] D.P.Bavirisetti, R.Dhuli, "Two-scale image fusion of visible and infrared images using saliency detection," Infrared Phys. Techn., 2016, 76: 52-64.

[18] J. Ma, Z. Zhou, B. Wang, H. Zong, "Infrared and visible image fusion based on visual saliency map and weighted least square optimization," Infrared Phys. Techn., 2017, 82: 8-17.

[19] https://figshare.com/articles/TNO_Image_Fusion_Dataset

[20] J.W.Roberts, J.Van Aardt, F.Ahmed, "Assessment of image fusion procedures using entropy, image quality, and multispectral classification," J. Appl. Remote Sens., 2008, 2(1) : 023522.

[21] Y.Han, Y.Cai, Y.Cao, X.Xu, "A new image fusion performance metric based on visual information fidelity," Information Fusion, 2013, 14(2): 127-135.

[22] L.Guo, X.Cao, L.Liu, "A novel autofocus measure based on weighted walsh-hadamard transform," IEEE Access, 2019, 7: 22107-22117.