Infrared and Visible Image Fusion Based on Gradient Reconstruction

Zhen Lei\textsuperscript{1,*}, Yanyan Huang\textsuperscript{2,*}, Peizhi Cui\textsuperscript{1,\textsuperscript{b}}

\textsuperscript{1}Army Armored Force Academy, Beijing 100072, China
\textsuperscript{2}Automation college, Nanjing University of Science and Technology, Nanjing 210094, China

\textsuperscript{*}Corresponding author e-mail: nudt_william_0809@163.com, \textsuperscript{*}yljx0322@yeah.net, \textsuperscript{b}zgy_0709zwjs@126.com

Abstract. A fusion method of infrared and visible image based on gradient reconstruction is proposed. Firstly, the structure tensor of infrared image and visible image is pre fused, and then the pre fused image is re fused based on gradient reconstruction and Poisson equation to get the final fusion image. The experimental results show that the method is effective, and the fusion image can synthesize better target feature information of infrared image and scene detail information of visible image, which has better performance than other image fusion methods.

1. Introduction
Visible imaging sensor and infrared imaging sensor are two kinds of commonly used sensors with different working mechanism. The infrared imaging sensor can identify the target by detecting the difference of thermal radiation between the target and the background, so it has a special ability of identification and camouflage. It can find people and various wheeled or tracked vehicles hidden in the forest and grass, and has a strong penetration ability. The image formed by it can still show the outline of the scene even though at night. But it’s insensitive to the brightness changes and has low imaging resolution. The visible light imaging sensor has a high definition, which can better reflect the details of the scene where the target is located, but it has nothing to do with the thermal contrast of the target scene. Under the bad weather conditions, the penetration ability to the atmosphere is poor, and the imaging ability at night is worse. For surveillance and reconnaissance mission, it is not only to find and identify the target, but also to determine the exact location of the target, which requires that the acquired image has a high definition. If the visible light and infrared image are organically fused, the image blur and disappearance caused by light, atmosphere, object occlusion, etc. can be eliminated, which is conducive to the integration of better target feature information of infrared image and scene detail information of visible light image to get a more comprehensive fusion image with more information and effectively improve the detection and identification ability of target [1-3].

Image fusion is a kind of information technology that can register the image data of the same object from different sources, and then combine the information advantages or complementarities contained in each image data organically with certain algorithm to generate new image data. Compared with a single information source, the new data can reduce or suppress the polysemy, incompleteness, uncertainty and...
error that may exist in the interpretation of the perceived object or environment, and make maximum use of the information provided by various information sources.

Commonly used image fusion methods include weighted average fusion method, HIS color mapping based fusion method, PCA based fusion method, pyramid decomposition based fusion method, wavelet transform based fusion method [4-6]. Among them, the fusion method based on HIS color mapping and the fusion method based on PCA are generally suitable for the fusion of multispectral image and panchromatic image, which will have some limitations when used for the fusion of visible light and infrared image. Although the weighted average fusion method can be used not only for the fusion of visible light and infrared image, but also for the fusion of multispectral image and panchromatic image, the difficulty lies in the determination of the weight value, and the method reduces the contrast of the image, which makes the edge and contour of the image fuzzy to a certain extent.

In this paper, the research of multisensor image fusion is made a bold attempt, and a multisensor image fusion algorithm based on gradient reconstruction is proposed. The gradient and structure information of the image are used to guide the fusion processing, and the iteration speed of Poisson equation is increased by using Gauss iteration instead of Jacobi iteration.

2. Image fusion based on gradient reconstruction

2.1. Image denoising based on adaptive Wiener filter

First, the output of the filter is adjusted according to the local variance of the infrared and visible image. The larger the local variance of the image is, the stronger the smoothing effect of the filter is. The filtering effect of this method is better than that of other traditional filters [7] (such as mean filter). It has obvious effect on the edge and other high-frequency parts of the reserved image, which is helpful for the following target detection or target recognition tasks.

2.2. Pre fusion based on structure tensor

Step 1: To calculate the gradient of infrared image and visible image respectively. The gradient of an image can be regarded as a two-dimensional discrete function. The gradient is the derivative of this two-dimensional discrete function. The calculation process is as follows

\[ [FX_1, FY_1] = \text{gradient}(F1) \]  
\[ FX_1 = \frac{\partial F_1}{\partial X}, FY_1 = \frac{\partial F_1}{\partial Y} \]  
\[ [FX_2, FY_2] = \text{gradient}(F2) \]  
\[ FX_2 = \frac{\partial F_2}{\partial X}, FY_2 = \frac{\partial F_2}{\partial Y} \]

Step 2: For each pixel, the structure tensor of the infrared image and the visible image is calculated. Structure tensor is a structure matrix of image, which is derived from the gradient of function. It summarizes the main direction of gradient in the specified neighborhood of a point and the degree of coherence in these directions, and can distinguish the flat area, edge area and corner area of image. The structure tensor of infrared image is

\[ ST_1 = \begin{bmatrix} FX_1^2 & FX_1 \cdot FY_1 \\ FX_1 \cdot FY_1 & FY_1^2 \end{bmatrix} \]  

The structure tensor of visible image is

\[ ST_2 = \begin{bmatrix} FX_2^2 & FX_2 \cdot FY_2 \\ FX_2 \cdot FY_2 & FY_2^2 \end{bmatrix} \]
ST2 = \[
\begin{bmatrix}
FX2^2 & FX2 \cdot FY2 \\
FX2 \cdot FY2 & FY2^2
\end{bmatrix}
\] (6)

Step 3: Sum the structure tensor of infrared image and visible image, get the matrix S_ST, get the diagonal matrix D composed of the eigenvalues of S_ST, and the full matrix V composed of the eigenvectors of S_ST.

Step 4: To make Q = 0.5 * \[\|FX1 + FX2\| + \|FY1 + FY2\|\], find the largest element D_{max} in the diagonal matrix D, then

\[FI = \text{sign}(V(:,2) \cdot Q) \cdot \sqrt{D_{\text{max}}} \cdot V(:,2)\] (7)

2.3. Secondary fusion based on gradient reconstruction

Step 1: Define the first-order backward difference operators K_x and K_y in horizontal and vertical directions as follows

\[K_x = \begin{bmatrix}
0 & 0 & 0 \\
-1 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix}, K_y = \begin{bmatrix}
0 & -1 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix}\] (8)

Step 2: The inner product of the x and y direction image blocks of the pre fused image is calculated by the difference operator, the sum is assigned to the current pixel of the output image, the first-order backward difference in the horizontal and vertical direction of the pre fused image is calculated, and the sum of the two is calculated.

Step 3: Use FFT to solve Poisson equation. The row and column Discrete Sine Transforms of the weighted sine inverse transform of the image to be recovered are calculated successively, and the image I_{rec} is reconstructed.

It should be noted that the Jacobi iteration process of Poisson equation is generally slow, and the result is not ideal after thousands of iterations. In order to speed up the iteration speed of Poisson equation, this paper adopts Gauss iteration. After the improvement of Gauss iteration algorithm, the result is better for solving Poisson equation with the same number of steps. In addition, it can be seen from the iterative solution process that with the increase of the value in the gradient domain, the contour of the image becomes clearer and clearer, otherwise, the image becomes more blurred. For the task of target detection, tracking and recognition, the clearer the target contour is, the more convenient the subsequent detection and recognition.

Step 4: Normalize the reconstruction image I_{rec} from 0 to 255. To find the maximum element I_{rec}^{\text{max}} and the minimum element I_{rec}^{\text{min}} in the reconstructed image I_{rec}, then

\[I_{rec} = 255 \cdot \frac{I_{rec} - I_{rec}^{\text{min}}}{I_{rec}^{\text{max}} - I_{rec}^{\text{min}}}\] (9)

2.4. Evaluation of fusion effect

An important step of image fusion is to evaluate the fusion effect, because human vision is not very sensitive to all kinds of changes in the image, and the visual quality of the image depends on the observer to a large extent, with subjectivity and incompleteness, so we should try to use the method of objective evaluation of fusion effect. Common objective evaluation methods include information entropy, mutual information \[8\].

(1) Information entropy. Information entropy is an important index to measure the richness of image information. Its expression is as follows
\[ H(X) = -\sum_{i=1}^{n} p_i \log p_i \] (10)

The size of information entropy reflects the amount of information carried by the image. The larger the entropy of the fusion result image is, the richer the information contained in the fusion result image is, the better the fusion effect is.

(2) Mutual information. Mutual information can be used to measure how much information the fusion image inherits from the source image. For two discrete random variables \( X \) and \( Y \), their mutual information is expressed as

\[ I(X,Y) = H(X) - H(X|Y) \] (11)

3. Experimental results and analysis

In order to verify the effectiveness of the proposed fusion method, two sets of real image sequences are selected for experiments. The visible and infrared image sequences to be fused are from TNO Institute of human factors in the Netherlands, three of which are taken respectively, as shown in Fig.1.

![Figure 1. Visible and infrared image sequences to be fused](image)

Fig.2 (a) is the result of each frame of visible light image and infrared image fusion by using the classical image fusion method based on gradient pyramid transform[5] (hereinafter referred to as algorithm 1), Fig.2 (b) is the result of each frame of visible light image and infrared image fusion by using the wavelet transform[6] (hereinafter referred to as algorithm 2), Fig.2 (c) is the result of each frame of visible light image and infrared image fusion by using the fusion method proposed in this paper (hereinafter referred to as algorithm 3).
Figure 2. Image sequence after fusion of different algorithms

Information entropy and mutual information are used for objective evaluation, and performance evaluation indexes of each fusion algorithm are shown in Table 1.

Table 1. Performance evaluation indexes of different image fusion algorithms

| Frame Number | 1       | 2       | 3       |
|--------------|---------|---------|---------|
| Information entropy |         |         |         |
| algorithm 1  | 6.3157  | 6.3118  | 6.3013  |
| algorithm 2  | 6.3680  | 6.3597  | 6.3547  |
| algorithm 3  | 6.9203  | 7.0948  | 7.1072  |
| Mutual information |       |         |         |
| algorithm 1  | 1.5295  | 1.5285  | 1.5189  |
| algorithm 2  | 1.5006  | 1.4905  | 1.4888  |
| algorithm 3  | 2.0164  | 2.5082  | 2.3756  |

It can be seen from table 1 that the information entropy and mutual information of the fusion image of the fusion algorithm in this paper are higher than those of other fusion methods, and the fusion image obtained is the clearest, which can better retain the target information of the infrared image, and has the background information with high definition, which is a feasible and effective fusion method.

4. Conclusions
In this paper, a method of infrared and visible image fusion based on gradient reconstruction is proposed. Firstly, the adaptive Wiener filter is used to denoise the image, then the structure tensor of infrared image and visible image is used to pre fuse, and finally the final image is obtained by the second fusion of the pre fusion image based on gradient reconstruction and Poisson equation. The experimental results show that the method is effective.

Acknowledgements
This work was financially supported in part by the National Natural Science Foundation of China (Grant No. 61374186).
References

[1] Bingwen C, Wenwei W, Qian-qing Q. Target detection in thermal-visible surveillance based on multiple-valued immune network. Journal of Infrared Millim Waves, 2014, 33(6): 654-659.

[2] Leykin A, Hammoud R. Pedestrian tracking by fusion of thermal-visible surveillance videos[J]. Machine Vision and Applications, 2010, 21(4): 587-595.

[3] Krotosky S J, Trivedi M M. Person surveillance using visual and infrared imagery[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18 (8) : 1096-1105.

[4] Xiaochuan Z. Sensor information fusion-matlab program implementation. China Machine Press, 2014, 127-151.

[5] Stoica I, Morris R, Karger D, et al. Cho rd: A scalable peer-to-peer lookup ser vice fo r inter net applications [A]. Processings of ACM SIGCOMM[C]. New York: ACM Press, 2001, 149-160.

[6] Freedman M, Mazieres D. Sloppy hashing and self-organize clusters[A]. Processings of IPTPS [C]. Heidelberg: Springer-Verlag, 2003, 45-55.

[7] Jin F, Fieguth P, Winger L, et.al. Adaptive Wiener filtering of noisy images and image Sequences. Processings of IEEE International Conference on Image Processing, 2003, 3:14-17.

[8] Byers J, Considine J, Mitzenmacher M. Simple load balancing in distributed hashing tables [A]. Processings of IPTPS [C]. Heidelberg : Springer-Verlag, 2003, 80-87.