Visible and Infrared Image Fusion Using Low-Frequency Coefficients Mapping Fusion Rule in the LWT

Xiaowei Wang [1,a,*], Guojun Lai[2,b], XuDong Wang[1,c], wenlong Hao[1,d]

[1] Research Center of Unmanned Aerial Vehicle, Army Aviation Institute, Beijing, China
[2] Departmet of Command, Army Aviation Institute, Beijing, China
[a] wangxiaoweis@163.com, [b] Laiguojun_lh@sina.com, [c] wanghx96@163.com, [d]574321214@qq.com

Keywords: Image Fusion, Lifting Wavelet Transform, Fusion Rules, Local Square Maximum, Weighted Average, Low Frequency Coefficients Mapping

Abstract: In order to improve the effect of detect object in visible and infrared fusion image, a Low-Frequency Coefficients Mapping fusion rule is proposed. The overall fusion scheme based on lifting wavelet transforms. Firstly, the source images of same scene are decomposed using lifting wavelet transform (LWT). Secondly, a Low-Frequency Coefficients Mapping fusion rule is used to select low frequency lifting wavelet coefficients of the visible and infrared images. The fusion rule of local square maximum is used to combine corresponding high frequency coefficients. After fusing low and high frequency coefficients of the source images, the final fused image is obtained using the inverse LWT. The experiments show that the proposed Low Frequency Coefficients Mapping Fusion Rule in the LWT obtains a good fusion results as compared to previous image fusion rule in the LWT such as local energy weighted average and average gradient maxima.

1. Introduction

Visible spectral and infrared spectral sensors are two kinds of the most commonly used sensors. The infrared imaging sensor is only sensitive to radiation of the target scene, which is mainly determined by its emissivity and temperature difference. And visible spectral imaging sensor is only sensitive to the reflection of target scene instead of thermal contrast. Through image fusion, images from visible spectral and infrared spectral sensors are processed in different aspects and levels to achieve more reliable information of the target. Image fusion is widely used in remote sensing, automatic target recognition, machine vision and other fields[1,2].

Many fusion techniques and algorithms for pixel-level fusion have been proposed, including the simplest weighted averaging, advanced pyramidal decomposition and wavelet transform methods. In this paper, a new image fusion rule based on LWT is put forward. The paper is organized as follows. In Section 2, the proposed LWT-based low frequency coefficients mapping image fusion

rule for visible and infrared images is described. The experiment results are given in Section 3. Finally, Section 4 provides the conclusion.

2. Proposed scheme

2.1. The Proposed Fusion Approach

The proposed method of visible and infrared image fusion uses the lifting wavelet transform for decomposition and reconstruction of the source images. The overall fusion scheme based on LWT is shown in Figure 1. Firstly we decompose source images of same scene using LWT and then low frequency coefficients are obtained using low frequency coefficients mapping fusion rule. The fusion rule of local square maximum is used to combine corresponding high frequency coefficients. After fusing low and high frequency coefficients of the source images, the final fused image is obtained using the inverse LWT. We strongly encourage authors to use this document for the preparation of the camera-ready. Please follow the instructions closely in order to make the volume look as uniform as possible.

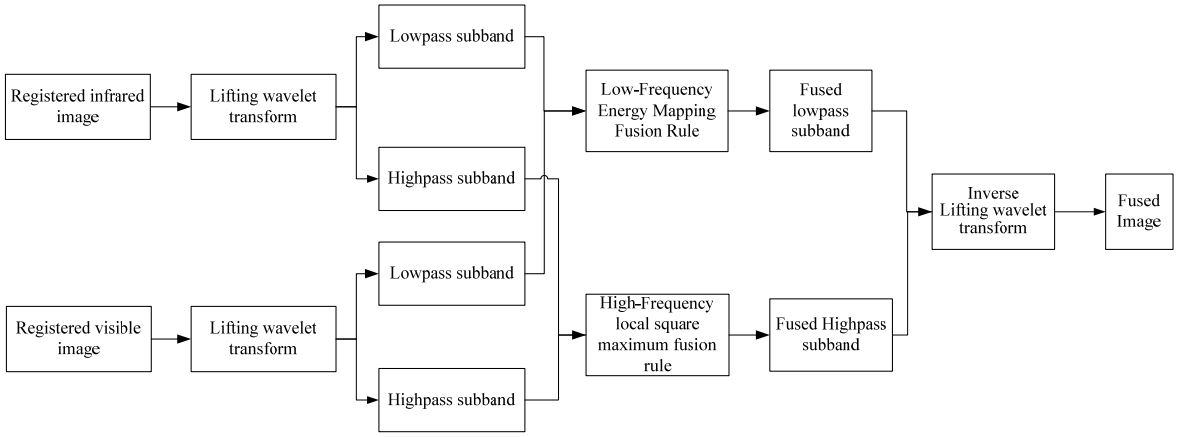


Figure 1 Block diagram of proposed system

2.2. The basic principle of lifting wavelet

Lifting scheme[3,4,5], put forward by Sweldens and Daubechies in the 1990s, is a kind of wavelet construction method does not rely on the Fourier transform. The wavelet transform based on lifting method is also called second generation wavelet transform. It can not only maintain the time-frequency localization and other characteristics of traditional wavelet, but also overcome limitations of its long time execution and the need of requiring a large memory. Lifting wavelet algorithm realization is divided into three steps: division, prediction, update and reconstruct.

1) Split

The original signal X(n) (n is behalf of the resolution) is divided into a low-resolution $x_{even}(n)$ (even-numbered sequence(2n)) and$x_{odd}(n)$ (odd-numbered sequence $X(2n+1)$) two parts in this step:

$$\begin{cases} x_{even}(n) = x(2n) \\ x_{odd}(n) = x(2n+1) \end{cases} \tag{1}$$

The signal is simply divided into two parts in this step, and does not change the description of the signal. The next lifting step is adopted to regroup both sequences to reduce the correlation between them.

2) Predict

Predict$x_{odd}(n)$ and define prediction P by $x_{even}(n)$. The prediction signal is as follows:

$$d(n) = x_{odd}(n) - P[x_{even}(n)] \tag{2}$$

The value of the error between $x_{odd}(n)$ and the predicted value $P[x_{even}(n)]$ represents the signal details $d(n)$.

c) Update

In order to maintain certain properties of the original signal X(n), update is defined and update with the subset of the data $d(n)$, as follows:

$$c(n) = x_{even}(n) + U[d(n)] \tag{3}$$

This step is known as the primal lifting in lifting method. If the decomposed approximate signal performs the above three decomposition steps and after a certain number of iterations, we can get a multi-level decomposition of the original signal. As can be seen from above, using lifting method to implement wavelet decomposition has the advantage that it can decompose wavelet transformation into a few very simple steps.

4) Reconstruction

Reconstruction is the inverse step of the decomposition process. It has three steps: anti-Update, anti- predict and merge. The reconstruction equation is as follows:

$$\begin{aligned} x_{even}(n) &= c(n) - U[d(n)] \\ x_{odd}(n) &= d(n) + P[x_{even}(n)] \\ x(n) &= \mathrm{Merge}[x_{even}(n), x_{odd}(n)] \end{aligned} \tag{4}$$

Image fusion algorithm based on lifting wavelet has following steps.

a)At First, take one-dimensional lifting transform by row direction in image matrixes of fused image A and B, get the approximate coefficient matrix and the details of the coefficient matrix respectively; the predict and update equation is as follows:

$$\begin{cases} d(n) = x_o(n) - x_e(n) \\ c(n) = x_e(n) + \frac{d(n)}{2} \end{cases} \tag{5}$$

b) Then take one-dimensional lifting transform by column-direction in the approximate coefficient matrix and the detail coefficients matrix that get from decomposition of A and B, get a low frequency coefficient matrix and three high-frequency coefficient matrix of the image, which is the completion of a layer wavelet decomposition of the image.

Repeat decomposition process to coefficient matrix of the image, the image can be wavelet decomposed at any scale as traditional.

c) Use certain fusion rules for low frequency coefficients and high frequency coefficients from the decomposition of image A, B, and then get fused image through inverse transform.

2.3. Fusion Rule

The low-frequency part reflects the overview of the source image at this resolution, including key details of the source image information. The high-frequency part contains a large number of the source edge details. Considering the characteristics of decomposition sub-bands, we adopt different fusion rules to low-frequency coefficient and high-frequency coefficient in this section.

2.3.1. Low-Frequency coefficient mapping Fusion Rule

To improve target's thermal feature, a low-frequency coefficients mapping fusion rule is proposed, in which the weights are obtained using the low frequency lifting wavelet coefficients mapping of the infrared image:

$$\mathrm{L_C^N(x,y)} = \alpha_1(\mathrm{x,y})\mathrm{L_A^N(x,y)} + \alpha_2(\mathrm{x,y})\mathrm{L_B^N(x,y)} \tag{6}$$

Where $\mathrm{L_A^N(x,y)}$, $\mathrm{L_B^N(x,y)}$expresses visible and infrared image's low-Frequency coefficient matrix at the N-th lifting wavelet decompose level. $\mathrm{L_C^N(x,y)}$ denotes the fused image's low-frequency coefficient matrix. The weights$\alpha_1(\mathrm{x,y})$, $\alpha_2(\mathrm{x,y})$are estimated as:

$$\alpha_2(\mathrm{x,y}) = 1 - \alpha_1(\mathrm{x,y})$$
$$\alpha_2(\mathrm{x,y}) = \begin{cases} \mathrm{k_1 \times \left[L_B^N(x,y) - E\right]} & \mathrm{L_B^N(x,y) - E \geq 0} \\ \mathrm{k_2 \times \left[E - L_B^N(x,y)\right]} & \mathrm{L_B^N(x,y) - E < 0} \end{cases} \tag{7}$$

Where E denoted low frequency coefficient mean of infrared image B. When the weight is zero, this means the substitution of an image by another. k_1, k_2 are estimated as:

$$\begin{cases} k_1 = \dfrac{1-b}{L_B^{\max} - E} \\ k_2 = \dfrac{1-b}{E - L_B^{\min}} \end{cases} \tag{8}$$

Where b denote the visible image's basic weight. The parameters $L_B^{\max}$, $L_B^{\min}$, denote the infrared image's maxima and minimum low-frequency coefficient respectively. The figure2 displays the proposed low-frequency coefficients mapping function curve. The main merits of this new low-frequency fusion rule are that it preserves the target temperature and visible features.

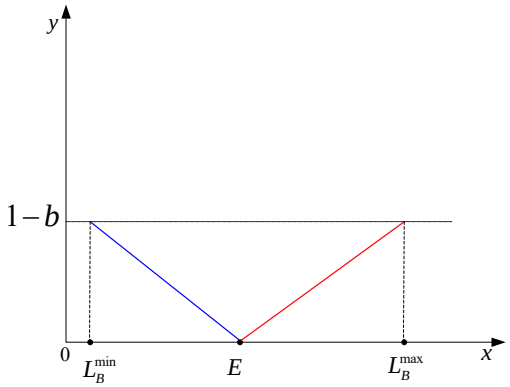


Figure 2 low-frequency coefficients mapping function curve

2.3.2. Fusion Rule of High-Frequency coefficient

We use the fusion rule of local square maximum to combine corresponding high sub-band coefficients. We first calculate the local square features $\mathrm{V_N^A(x,y)}$ and $\mathrm{V_N^B(x,y)}$ of the N-th high-frequency coefficients of the visible and infrared images. $\mathrm{V_N^A(x,y)}$,$\mathrm{(x,y)}$are calculated over a 3-by-3 or 5-by-5 window using the formula:

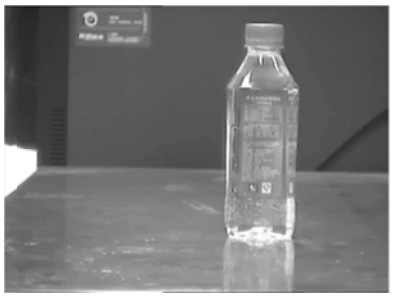
$$\begin{cases} V_N^A(x,y) = \sum_{M=-1}^{M=1} \sum_{N=-1}^{N} \left[H_N^A(x+M, y+N) - \mu_A \right]^2 \\ V_N^B(x,y) = \sum_{M=-1}^{M=1} \sum_{N=-1}^{N} \left[H_N^B(x+M, y+N) - \mu_B \right]^2 \end{cases} \tag{9}$$

Where $H_N^A(x,y)$, $H_N^B(x,y)$ expresses visible and infrared image's high-Frequency coefficient matrix at the N-th lifting wavelet decomposes level. The parameterμ_A, $\mu_B$denoted high frequency coefficient region mean of visible and infrared image respectively. Then, the high-frequency coefficient decision map is:

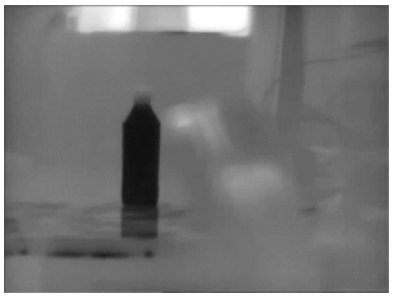
$$F_N(x,y) = \begin{cases} V_N^A(x,y) & \text{if } V_N^A(x,y) \geq V_N^B(x,y) \\ V_N^B(x,y) & \text{otherwise} \end{cases} \tag{10}$$

3. Experiment result

To evaluate the performance of the proposed fusion method, several experimental results are presented in this section. Figure 3 corresponds to the 'bottle' frame pair and its fusion result whose size is 320×240. And Figure 4 corresponds to the 'UN camp 'frame pair and its fusion result with the size of 320x240. The experiment was implemented on Intel® Core™ i7-4690 CPU @3.60Hz, 4.00GB PC platform equipped with visual c + + 6.0 simulating environment. Figure 3(c), 2(d) and 3(e) displays the final fused image respectively by using three fusion algorithms, it contains more information than both of Figure 3(a) and Figure 3(b). Algorithm 1 uses Low-frequency Coefficients Mapping and high-frequency local square maximal value fusion rule in the LWT. Algorithm 2 uses Low-frequency local energy weighted average and high-frequency local square maximal value fusion rule. Algorithm 3 uses Low-frequency Coefficients Mapping and high-frequency average gradient fusion rule.



(a) The visible image



(b) The infrared image

(c) The fused image using algorithm1

(d) The fused image using algorithm2

(e) The fused image using algorithm3

Figure 3 Fusion results of methods

To evaluate the effectiveness of the fusion algorithm, the paper adopts following principals: evaluating the fusion result through five indexes information entropy, average gradient, standard deviation, and Spatial frequency, Speed, which do not require ground truth images for evaluation. Table 1 is the comparison of image fusion results. From the table 1we can see that the information entropy, standard deviation and spatial frequency gained by algorithm1 proposed in this paper are highest respectively, and only average gradient is decreased by 21.24% than algorithm 3, but the proposed algorithm expense little time. The results presented in this example can demonstrate that our approach can fuse the visible and infrared images while retaining much more information than that of the other two methods.

Table 1 the performance of different fusion methods using different statistical criteria

statistical criteria / method	entropy	Average Gradient	Standard deviation	Spatial frequency	Speed (frames/s)
Infrared image	5.49063	1.35639	32.2235	4.12311	
Visible image	6.77162	4.6404	38.1527	12.0416	
Algorithm 1	6.64062	3.62055	30.9127	10.63071	125
Algorithm 2	6.27548	3.18235	24.6223	11.4018	121
Algorithm 3	6.6327	4.59737	30.8554	10.6301	61

From Figure 4(a) we can see that in visible spectral images the environment backgrounds are clear, while the human targets are too dark to recognize. On the other hand, Figure 4(b) show infrared source images in which the human targets are salient, while backgrounds are blur. Figure 4(c) , 4(d) and 4(e) show the fusion results by using three fusion algorithm

(a) the visible source image

(b) the Infrared source image

(c) The fused image using algorithm1

(d) The fused image using algorithm2

(e) The fused image using algorithm3

Figure 4 Fusion results of methods

Table2 The performance of different fusion methods using different statistical criteria

statistical criteria / method	entropy	Average Gradient	Standard deviation	Spatial frequency	Speed (frames/s)
Infrared image	5.49063	1.35639	32.2235	4.12311	
Visible image	6.77162	4.6404	38.1527	12.0416	
Algorithm 1	7.05835	7.09185	39.0863	11.4018	119
Algorithm 2	6.47013	6.35567	27.7883	10.6301	126
Algorithm 3	7.05474	6.9773	39.0406	11.4018	39

From Table 2, it can be seen the fusion image based on the proposed method is the highest on Entropy, Average Gradient, Standard deviation, spatial frequency and Speed. It indicates that the featured information and average information volume are comparatively highest. Meanwhile, it proves that the image is the clearest.

4. Conclusions

In this paper, by using LWT, we have decomposed source visible and infrared image into couples of sub-band images which are separately approximate images or detail images. The low frequency components give an overview of the target object or scene and the high frequency components highlight the detail information. By using low-frequency coefficients mapping and high- frequency coefficients local square maximum fusion rules to sub-band images, we can make full use of the temperature and visible features on the target object or scene. Experimental results show that the proposed image fusion rule obtained better results in LWT domain when tested with performance measures, Entropy, Average Gradient, Standard deviation, Spatial frequency and Speed.

Acknowledgements

Thank this template provide AIMS. This paper is completed with the help of my colleagues. Thank them who previously provided technical support.

References

[1] J. Li, Y. Jiang, R. Fan. Recognition of Biological Signal Mixed Based on Wavelet Analysis[C]. In: Y. Jiang, et al (eds). Proc. of UK-China Sports Engineering Workshop. Liverpool: World Academic Union. 2007: 1~8.

[2] Sweldens W(1997). The Lifting Scheme: A Construction of Second Generation Wavelets [J]. SIAM J Math Anal, Vol. 29(2):511~546.

[3] Suwen Zhang, Pei Xiong, Juan Chen. An Image Fusion Method Based on Lifting Wavelet and Weighted Nonnegative Matrix Factorization[C].2010 International Conference on Measuring Technology and Mechatronics Automation: 484~487.

[4] Xu Qiang, Cheng Yinglei, Zhao Huizhen. A Fast Fusion Approach of Remote Sensing Images Based on Lifting Wavelet[C].2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC), Dec 20-22, 2013, Shenyang, China,1065~1070.

[5] YIN Dehui, LI Bingfa, TANG Yan. Fusion Algorithm Based on Wavelet Transform [J]. Journal of System Simulation. 2006, 18(5):1289~1291.