Visible and Infrared Image Fusion using the Lifting Wavelet

Yuelin Zou1*, Xiaoqiang Liang1, Tongming Wang2
1Software engineering department, Shijiazhuang Information Engineering Vocational College, Shijiazhuang 050035, PR China
2The Center of Education Technology, Hengshui University, Hengshui, 053000, PR China
*Corresponding author, e-mail: 53823953@qq.com

Abstract

In recent years image fusion plays a vital role in the image processing area. Fused images would help in doing many applications in image processing like segmentation, image enhancement and many. In order to improve the effect of fusion visible and infrared image images of the same scene, this paper presents an image fusion method based on lifting wavelet domain. Firstly, the source images are decomposed using lifting wavelet domain transform (LWT). Secondly, a weighted average approach based on normalized Shannon entropy is used to fuse low frequency lifting wavelet coefficients of the visible and infrared images. The fusion rule of local energy 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 method provides improved subjective and objectives results as compared to previous image fusion methods such as Laplacian transform and traditional Wavelet transform.

Keywords: LWT, LEM, image fusion, weighted average

Copyright © 2013 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

Visible images are the images obtained in the visible spectrum and vary according to the illumination conditions. Visible image holds the details of the necessary features of the images, required in the process of biometric authentication [1]. Since visible images vary according to the luminance and conditions under which the test images are taken, they are liable to result in error or otherwise wrong recognition. Infrared images are captured using an infrared sensor camera in the far infrared region. Infrared Image gives the measure of energy radiations from the object and it is less sensitive to illumination changes and is even operable in darkness [2]. The energy radiated from the object, may change according to the surroundings and due to the physical exertions. The features of the object, the primary requisite for acquiring the correlation with the database images are indistinguishable in case of Infrared image [3]. In addition, infrared image as a standalone does not provide high-resolution data [4] Hence, fusion of visible and infrared images provides better solution to achieve the best feature of both the images for target recognition system [5]. Hence this paper proposes a novel scheme of fusion visible and infrared images.

Because of the multi-resolution and good properties of time–frequency analysis, wavelet transform [6] reveals its good performance in the image fusion field. Singh proposed a fusion scheme based on pixel, employs genetic algorithms to operate the traditional wavelet coefficients to decide how to combine IR with visible information [7]. However, the genetic algorithms require a large amount of the time to work. Hariharan proposed a new image fusion technique, utilizing Empirical Mode Decomposition (EMD), for improved face recognition. EMD is a non-parametric data-driven analysis tool that decomposes non-linear non-stationary signals into Intrinsic Mode Functions (IMFs) [8]. We can know the EMD has High time complexity. A fusion algorithm is proposed to combine pairs of multispectral magnetic resonance imaging such as T1, T2 and proton density brain images. The proposed algorithm utilizes different features of redundant discrete wavelet transform, mutual information based non-linear registration and entropy information to improve performance [7]. However, it can only capture limited orientation information, including horizontal, vertical and diagonal orientations in each decomposition stage [9]. In order to overcome the problem, lifting wavelet transform is proposed to use to fuse images. The normalized Shannon entropy is adopted as the fusion rule in low frequency coefficient. The local energy feature is used to select the fusion coefficient in high-Frequency coefficient. The experiments demonstrate the fusion method is effective to fuse the visible and infrared image fusion and outperform the DWT based method and Laplacian based method.
2. The Proposed Method

2.1. Lifting Wavelet Transform

Lifting scheme, put forward by Swelden in the 1990s, is a kind of wavelet construction method does not rely on the Fourier transform. Compared with traditional WT (wavelet transform, WT), LWT (Lifting Wavelet Transform) possesses several advantages, including possibility of adaptive and nonlinear design, in place calculations, irregular samples and integral transform. It can be seen as an alternate implementation of traditional wavelet transform. The main feature of the lifting wavelet transform is that it provides an entirely spatial domain interpretation of the transform, in contrast to the traditional frequency domain based constructions [10] lifting wavelet algorithm realization is divided into three steps: division, prediction and update. \( P \) and \( U \) denote the prediction and update operator of the lifting wavelet at level \( l \), respectively. \( a \) is \( l+1 \) level decomposition by LWT of the input signal. \( d_{l+1} \) and \( a_{l+1} \) respectively are the detail and approximate coefficients after LWT decomposition of the \( a_{l} \). The prediction coefficient and update coefficients at level \( l \) in the lifting wavelet transform are expressed as follows:

\[
p^{l+1} = p_0, 0, \ldots, p_{2^l-2}, 0, \ldots, p_{m-1}\tag{1}
\]

\[
u^{l+1} = u_0, 0, \ldots, 2^{l-1} u_{2^l-2}, 0, \ldots, 2^{l-1} u_{m-1}\tag{2}
\]

The decomposition results of an approximation signal at level via lifting stationary wavelet are expressed by following equations:

\[
d_{i+1} = a_i - p^{i+1}d_{i+1}, a_{i+1} = a_i + U^{i+1}d_{i+1}\tag{3}
\]

Where \( d_{i,1} \) and \( a_{i,1} \) are detail signal and approximation signal of \( a_i \) at level \( i+1 \).

The reconstruction procedure of lifting stationary wavelet transform is directly achieved from its forward transform, which is expressed as below.

\[
a_i = \frac{1}{2} \left( a_{i+1} - U^{i+1}d_{i+1} + d_{i+1} + P^{i+1}(a_{i+1} - U^{i+1}d_{i+1}) \right)\tag{4}
\]

The forward and inverse transform of lifting stationary wavelet transform is shown in Figure 1. Compared with the DWT, lifting wavelet transform do not downsample and upsample the highpass and the lowpass coefficients during the decomposition and reconstruction of the image. When LWT is introduced to image fusion, more information for fusion can be obtained and the impacts of mis-registration on the fused results can also be reduced effectively.

![Figure 1. Decomposition and Reconstruction Diagram of LWT](image)

2.2. Fusion Rule

Considering the characteristics of decomposition subbands, we adopt different fusion rules to low-frequency coefficient and high-frequency coefficient.

2.2.1. Fusion Rule of Low-Frequency Coefficient

An approach [11] is a weighted averaging proposed, in which the weights are obtained using a region-based activity measurement of the low frequency lifting wavelet coefficients:

\[
W_s(R) = \frac{1}{|R|} \sum_{n \in R} x_s^2(n) \times \log_2 x_s^2(n)\tag{5}
\]

Where \( R \) is the region with size \( |R| \), and \( x_s \) are the input low-Frequency wavelet coefficients, and \( L \) represents the coarsest resolution level. Hence, the composite low-Frequency coefficient is generated using:

Visible and Infrared Image Fusion using the Lifting Wavelet (Yuelin Zou)
Where \( x_u^L(i, j) \) and \( x_u^H(i, j) \) are the fused and input low-frequency coefficient wavelet coefficients, \( w_a(i, j) \) and \( w_b(i, j) \) are obtained using (6), and \( L \) represents the coarsest resolution level.

### 2.2.2. Fusion Rule of High-Frequency coefficient

We used the fusion rule of local energy maximum to combine corresponding subband coefficients. We first calculate the local energy features \( E_{A,i}^l \) and \( E_{B,i}^l \) of the high-frequency coefficient of the visible and infrared images. The local energy feature is defined as:

\[
E_{i,j}(x, y) = \sqrt{\sum_{x+p, y+q} W(x', y') \left[ D_{i,j}(x+x', y+y') \right]^2}
\]

(7)

Where \( D_{i,j}(x+x', y+y') \) denotes the decomposition coefficient at the \( i \)-th directional subband of the level \( l \). \( W \) is the kernel operator, the size of the local window is \( p \times q \), the center kernel operator coefficient \( W(0,0) \) is equal to \( 1/2 \), and the other kernel operator coefficients are equal to \( 1/[2 \times (p \times q - 1)] \). The size of the local energy feature window that we choose is \( 3 \times 3 \), so the \( W \) is:

\[
W = \begin{bmatrix}
1/16 & 1/16 & 1/16 \\
1/16 & 1/2 & 1/16 \\
1/16 & 1/16 & 1/16
\end{bmatrix}
\]

Then, the high-frequency coefficient decision map is:

\[
D_{i,j}(x, y) = \begin{cases} 
1 & \text{if } E_{A,i}^l(x, y) \geq E_{B,i}^l(x, y) \\
0 & \text{otherwise}
\end{cases}
\]

(8)

In order to keep the consistency in the high-frequency detailed components of the fused image, we adopt the “majority” principle to do consistency detection and modulation for fused decision map \( D_{i,j}(x, y) \). We can get the high frequency coefficients use the \( D_{i,j}(x, y) \) modified by the majority consistency detection.

### 2.3. Fusion Approach

We present the proposed visible and infrared image fusion algorithm. Figure 1 shows the block diagram of the proposed method, which consists of a number of essential stages:

1. **Step 1**: The source images are decomposed into different directions and scales using the LWT.
2. **Step 2**: Low frequency lifting wavelet coefficients of the final fused image are obtained via weighted averaging, in which the weights are obtained using (5).
3. **Step 3**: High frequency lifting wavelet coefficients of the source images are integrated using local energy maximum rule. “Majority” principle is adopted to do consistency detection and modulation for fused high-frequency coefficients.
4. **Step 4**: The inverse LWT of the new low and high frequency lifting wavelet coefficients generates the final fused image.

### 3. Results and Discussion

Two sets of images with perfect registration and one set of images with mis-registration are used to evaluate the proposed fusion algorithm. The proposed image fusion method was tested against several state-of-the-art image fusion methods including the simple averaging, the Laplacian Transform (LT), the discrete wavelet transform with the same fusion rules weighted average (WA) and local energy maximum (LEM) algorithm. Furthermore, proposed fusion method is compared with regional energy contrast pyramid algorithm with respect to the human visual characteristics \([12]\) called as Contrast Pyramid in this paper. In the LT method five decomposition levels is used for image decomposition. For the DWT based methods, the
available “db2” wavelet is used and five decomposition levels are used for image decomposition. In the Contrast Pyramid method, all the parameters used in the paper [12] is adopted.

In the fused image obtained by the proposed method, which is also implemented in the lifting wavelet transform domain, some details such as the contours of trees and the bright points are transferred into the fused image. In Figure 3, the roads’ details from the visible image and the person’s details from the infrared image, and also in Figure 4 the car’s details from the visible image and the house’s details from the infrared image, are better transferred into the fused image in the proposed method. Generally, it can be seen in Figure 3-4 that the person is brighter in the proposed method, and the contrast of the fused images is far better compared to other methods.

Figure 2. Schematic Diagram of LWT-based Fusion Algorithm

Figure 3. The Fusion Results of ‘Forest’ Image by Different Method
Two metrics are considered in this paper, which do not require ground truth images for evaluation. The first metric is $Q^{AB/F}$ [13], which considers the amount of edge information transferred from the input images to the fused image using a Sobel edge detector to calculate the strength and orientation information at each pixel in both source and the fused images. The second metric is the entropy index, which measures the information content in an image. An image with high information content will have high entropy. The third metric is the mutual information (MI) metric [14] used to evaluate the fusion performance quantitatively in this paper. Table 1 shows the average performance results from different image fusion methods and different datasets. 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.

4. Conclusion

In this paper, we have presented a new lifting wavelet based visible and infrared images fusion method. Proposing new fusion rules for merging high and low frequency wavelet coefficients, which is the second step in the wavelet-based image fusion, is the main novelty of this paper. The weighted average method and local energy maximum are respectively used on the low frequency and high frequency coefficients. The experimental results demonstrated that the proposed method outperforms the standard fusion methods in the fusion of infrared and visible images. The proposed image fusion algorithm is an effective, efficient and feasible algorithm. Finally, it is important to note that the proposed LWT-based fusion algorithm outperforms the DWT-based fusion algorithm in some cases.
Acknowledgement
Some of the images used in this paper are available in http://www.imagefusion.org. This work was supported in part by University Science Research Project of Hebei Province under grants 2011412 and Shijiazhuang Tiedao University project of Education and teaching reform under grants 110404. The authors also thank the anonymous referees for their valuable suggestions.

References