A SPATIAL DOMAIN AND FREQUENCY DOMAIN INTEGRATED APPROACH TO FUSION MULTIFOCUS IMAGES

J. Yang, Y. Ma, W. Yao, W. T. Lu

Institute of atmospheric sounding, Chinese Academy of Meteorological Sciences, Beijing, P. R. China
- yangjun100@gmail.com, maying@cams.cma.gov.cn, yaowen@cams.cma.gov.cn, wtlu@ustc.edu

KEY WORDS: multifocus, image fusion, Sum-modified-Laplacian, stationary wavelet transform

ABSTRACT:

A spatial domain and frequency domain integrated approach was proposed to fusion multifocus images. The proposed multifocus image fusion algorithm was composed of computing Sum-modified-Laplacian(SML) for each focus image, stationary wavelet transform(SWT) decomposition, image fusion and inverse SWT. Firstly, two initial binary decision maps were created by setting two thresholds to the SML difference between two focus images. Secondly, two different focus images were decomposed using SWT transform separately, then in the SWT domain of the two transformed images, the new SWT coefficients were acquired by adopting a simple fusion rule. Low-bands coefficients were integrated using the weighted average, and high-bands coefficients were integrated using choose max and the two SML maps. Finally the fused image was obtained by performing an inverse SWT transform. Two groups of different focus images were performed to evaluate the performance of our method. Experimental results showed that our algorithm can provide better performance than the wavelet method from both visual perception and quantitative analysis.

1. INTRODUCTION

Due to the limited depth-of-focus of optical lenses, it is often difficult to get an image that contains all relevant objects in focus. Multifocus image fusion methods are developed to solve this question. There are various approaches have been performed in the literatures. These approaches can be divided into two types, spatial domain method and frequency domain method.

Spatial domain fusion method is performed directly on the source images. Weighted average is the simplest spatial domain method, which needn't any transformation or decomposition on the original images. The merit of this method is simple and fit for real-time processing, but simple addition will reduce the signal-to-noise of the result image. Improved method is to compute the degree of focus for each pixel use various focus measures in multifocus images. A focus measure is defined which is a maximum for the best focused image and it generally decreases as the defocus increases (Krotkov,1987). Many focus measures techniques have been implemented in literatures, such as gray level variance(GLV) (Tyan,1997), Energy of image gradient(EOG), Energy of Laplacian of the image(EOL) (Huang and Jing, 2007), Sum-modified-Laplacian(SML) (Nayar and Nakagawa, 1994), Tenenbaum's algorithm(Tenengrad) and so on. Pixels with maximum focus measures are selected to construct and form ultimate all-in-focus image. The most commonly reported problem in this technique is from blocking effects.

In frequency domain methods, the input images are decomposed into multiscale coefficients initially. Various fusion rules are used in the selection or manipulation of these coefficients and synthesized via inverse transforms to form the fused image. Both pyramid and wavelet transforms are used as multiresolution filters. This type method can avoid blocking effects. However, many of these approaches, such as discrete wavelet transform(DWT) (Li et al.,1995), wavelet packet transform(WPT) (Yang and Zhao, 2007a) and Curvelet transform (Yang and Zhao,2007b), are shift-variant. So if there is a movement of the object in the source images or there is a misregistration of the source images, the performance of those algorithms will deteriorate. Some shift-invariant transforms are used to alleviate this phenomena including discrete wavelet frame transform(DWFT) (Li et al.,2002), dual tree complex wavelet transform(DT-CWT) (Ioannidou and Karathanassi,2007) and stationary wavelet transform(SWT) (Wang et al.,2003). Although these shift-invariant transforms are adopted, ringing effects have still been widely reported.

In order to overcome the disadvantages of the spatial domain method and the frequency domain method, we proposed a spatial domain and frequency domain integrated approach to fusion multifocus images in this paper. The simulation experiments obtained satisfactory results.

The next sections of this paper were organized as follows. In section 2 we provided a detail description of the Sum-modified-Laplacian and stationary wavelet transform. Section 3 presented our image fusion scheme. In section 4 two different focus images were used to evaluate our fusion algorithm. In the end, a conclusion was drawn in section 5.

2. SUM-MODIFIED-LAPLACIAN AND STATIONARY WAVELET TRANSFORM

1.1 2.1 Sum-modified-Laplacian

A focus measure is defined which is a maximum for the best focused image and it generally decreases as the defocus increases. Therefore, in the field of multifocus image fusion, the focused image areas of the source images must produce maximum focus measures, the defocused areas must produce minimum focus measures in contrast. Let $f(x, y)$ be the gray level intensity of pixel (x, y).

Nayar (1994) noted that in the case of the Laplacian the second derivatives in $x-$ and $y-$ directions can have opposite signs and tend to cancel each other. Therefore, he proposed the modified Laplacian (ML). The expression for the discrete approximation of ML is:

$$\nabla^2_{ML} f(x,y) = |2f(x,y) - f(x-step,y) - f(x+step,y)| + |2f(x,y) - f(x,y-step) - f(x,y+step)| \quad (1)$$

In order to accommodate for possible variations in the size of texture elements, Nayar (1994) used a variable spacing (step) between the pixels to compute ML. In this paper 'step' always equals to 1.

$$SML = \sum_{i=x-N}^{x+N} \sum_{j=y-N}^{y+N} \nabla^2_{ML} f(i,j) \quad \text{for} \quad \nabla^2_{ML} f(i,j) \ge T \quad (2)$$

where T is a discrimination threshold value. The parameter N determines the window size used to compute the focus measure.

$$\phi_{j,k}(x) = 2^{-j}\phi(2^{-j}x - k) \quad (4)$$

where $\phi(x)$ is the scaling function, which is a low-pass filter. $c_{j,k}$ is also called a discrete approximation at the resolution 2^j.

If $\varphi(x)$ is the wavelet function, the wavelet coefficients are obtained by

$$\omega_{j,k} = \left\langle f(x), 2^{-j}\varphi(2^{-j}x - k) \right\rangle \quad (5)$$

$\omega_{j,k}$ is called the discrete detail signal at the resolution 2^j.
As the scaling function $\phi(x)$ has the following property:

$$\frac{1}{2}\phi(\frac{x}{2}) = \sum_n h(n)\phi(x-n) \quad (6)$$

$c_{j+1,k}$ can be obtained by direct computation from $c_{j,k}$

$c_{j+1,k} = \sum_n h(n-2k)c_{j,n}$ and

$$\frac{1}{2}\varphi(\frac{x}{2}) = \sum_n g(n)\phi(x-n) \quad (7)$$

The scalar products $\left\langle f(x), 2^{-(j+1)}\varphi(2^{-(j+1)}x - k) \right\rangle$ are computed with

1.2 2.2 Stationary wavelet transform

In this section, we present the basic principles of the SWT method. In summary, the SWT method can be described as follows (Wang et al.,2003).

When the high pass and low pass filters are applied to the data at each level, the two new sequences have the same length as the original sequence without decimation. That is different from DWT, where decimation is necessary.

Supposing a function $f(x)$ is projected at each step j on the subset $V_j (LL \subset V_3 \subset V_2 \subset V_1 \subset V_0)$. This projection is defined by the scalar product $c_{j,k}$ of $f(x)$ with the scaling function $\phi(x)$ which is dilated and translated

$$c_{j,k} = \left\langle f(x), \phi_{j,k}(x) \right\rangle \quad (3)$$

$$\omega_{j+1,k} = \sum_n g(n-2k)c_{j,n} \quad (8)$$

Equations (7) and (8) are the multiresolution algorithm of the traditional DWT. In this algorithm, a downsampling procedure is performed after the filtering process. That is, one point out of two is kept during transformation. Therefore, the whole length of the function $f(x)$ will reduce by half after the transformation. This process continues until the length of the function becomes one.

However, for stationary or redundant wavelet transform, instead of downsampling, an upsampling procedure is carried out before performing convolution at each scale. The distance between samples increases by a factor of two from scale j to the next. $c_{j+1,k}$ is obtained by

$$c_{j+1,k} = \sum_l h(l)c_{j,k+2^j l} \quad (9)$$

and the discrete wavelet coefficients by

$$\omega_{j+1,k} = \sum_l g(l)c_{j,k+2^j l} \quad (10)$$

The redundancy of this transform facilitates the identification of salient features in a signal, especially for recognizing the noises.

This is the transform for one-dimensional signal. For a two dimensional image, we separate the variables x and y and have the following wavelets.

— Vertical wavelet: $\varphi^1(x,y) = \phi(x)\varphi(y)$
— Horizontal wavelet: $\varphi^2(x,y) = \varphi(x)\phi(y)$
— Diagonal wavelet: $\varphi^3(x,y) = \varphi(x)\varphi(y)$

Thus, the detail signals are contained in three subimages,

$$\omega_{j+1}^{1}(k_x,k_y)=\sum_{l_x=-\infty}^{+\infty}\sum_{l_y=-\infty}^{+\infty}g(l_x)h(l_y)c_{j,k+2^j}(l_x,l_y) \quad (11)$$

$$\omega_{j+1}^{2}(k_x,k_y)=\sum_{l_x=-\infty}^{+\infty}\sum_{l_y=-\infty}^{+\infty}h(l_x)g(l_y)c_{j,k+2^j}(l_x,l_y) \quad (12)$$

$$\omega_{j+1}^{3}(k_x,k_y)=\sum_{l_x=-\infty}^{+\infty}\sum_{l_y=-\infty}^{+\infty}g(l_x)g(l_y)c_{j,k+2^j}(l_x,l_y) \quad (13)$$

3. OUR PROPOSED FUSION ALGORITHM

An important preprocessing step in image fusion is image registration. It ensures that the information from each of the images refers to the same physical structure in the environment. In this paper, we assume that images to be combined have already been co-registered. The proposed multifocus image fusion algorithm is composed of computing SML for each focus image, SWT decomposition, image fusion and inverse SWT.

Firstly, we choose SML as focus measure to compute the clarity of each focus image. With the SML, we can get two initial binary decision maps by setting two thresholds to the SML difference between two focus images, which can be represented with the following equations:

$$Map1(x,y)=\begin{cases}1 & SML1(x,y)-SML2(x,y)>T1\\ 0 & SML1(x,y)-SML2(x,y)\le T1\end{cases} \quad (14)$$

$$Map2(x,y)=\begin{cases}1 & SML1(x,y)-SML2(x,y)<T2\\ 0 & SML1(x,y)-SML2(x,y)\ge T2\end{cases} \quad (15)$$

where $Map1$ and $Map2$ denote two decision maps. $SML1$ and $SML2$ represent the SML values of two focused images respectively. $T1$ and $T2$ are two thresholds.

Secondly, two focus images are decomposed into multiscale coefficients with SWT respectively. Due to the decomposition using SWT, the original image is transformed into four pieces which can be labeled as LL, LH, HL and HH. The LL piece is the low-band or called approximated component, the remaining pieces LH, HL and HH are high-bands or called detail components. All of the components have the same size as the original image due to shift-invariant character.

In the wavelet domains of the two transformed images, low-bands coefficients are integrated using the weighted average, the fusion equation is as below:

$$LL(x,y)=a\times LL1(x,y)+b\times LL2(x,y) \quad (16)$$

where LL represents the new low-band coefficient after fusion. a and b denote weighted coefficients, their summation is always 1.
The high-bands coefficients are first integrated using choose-max as follows:

$$HH(x,y)=\begin{cases}HH1(x,y) & HH1(x,y)>HH2(x,y)\\ HH2(x,y) & HH1(x,y)\le HH2(x,y)\end{cases} \quad (17)$$

Then the two SML decision maps are used to refine the fusion rule.

$$HH(x,y)=\begin{cases}HH1(x,y) & Map1(x,y)=1\\ HH2(x,y) & Map2(x,y)=1\\ HH(x,y) & others\end{cases} \quad (18)$$

The similar fusion rules are performed on LH and HL high-bands in each decomposition level.
At last, the fused image will be obtained by reconstructed with the fused approximate coefficients and detailed coefficients.

4. EXPERIMENTAL RESULT AND EVALUATION

To illustrate the performance of the proposed method, two groups of different focus but co-registered images are taken as examples in this paper. In order to compare fusion effect, discrete wavelet fusion method is performed as reference.
The proposed approach is implemented in personal computers with MATLAB 6.5 programs under Microsoft Windows XP environment.

The wavelet function sym4 is adopted and the input images are decomposed to 2 levels in this paper. The thresholds $T1$ and $T2$ are set to 0.2 and -0.2 respectively. Both low-band weighted coefficients a and b are equal to 0.5. The simulation experiment results are shown as Figure 1 and Figure 2.

It is difficult to evaluation the quality of a fusion image (Wald et al.,1997). Generally, the visual perception and quantitative analysis are used to compare image quality. From the visual perception, it is obvious form Figure 1 and Figure 2 that the proposed method has reserved more detail information than the wavelet transform method. The average gradient of image is computed to evaluation image quality quantitatively (Wang et al.,2002). The bigger of the average gradient, the more clear-cut of the image is. The equation of the average gradient is as follows,

$$\bar{g}=\frac{1}{(M-1)(N-1)}\cdot\sum_{i=1}^{M-1}\sum_{j=1}^{N-1}\sqrt{\frac{1}{2}\{[\frac{\partial f(x_i,y_j)}{\partial x_i}]^2+[\frac{\partial f(x_i,y_j)}{\partial y_j}]^2\}} \quad (19)$$


(a) Left focus image


(b) Right focus image


(c) Wavelet algorithm image


(d) Our algorithm image

Figure 1 First fusion experiment of multifocus image


(a) Left focus image


(b) Right focus image


(c) Wavelet algorithm image


(d) Our algorithm image

Figure 2 Second fusion experiment of multifocus image

We compute the average gradients of the two groups of images, the results are shows as Table 1.

	Left focus	Right focus	Wavelet	Our method
Experiment 1	2.5166	3.8978	4.2799	5.0357
Experiment 2	2.6161	2.8508	2.4947	3.2018

Table 1 Comparison of image average gradient

From Table 1, we can notice that our algorithm has higher average gradient than wavelet method and the defocused images, which demonstrators that our algorithm is valid and performs well.

5. CONCLUSION

There are two types of multifocus image fusion methods, the spatial domain and the frequency domain. However, they both have respective disadvantages. This paper proposed a spatial domain and frequency domain integrated approach to fusion multifocus image. Two groups of different focus images were performed to evaluate the performance of our method. Experimental results showed that our algorithm can provide better performance than the wavelet method from both visual perception and quantitative analysis.

REFERENCES

Huang, W., Jing, Z.,2007. Evaluation of focus measures in multi-focus image fusion. *Pattern Recognition Letters*, 28, p. 493-500.

Ioannidou, S.; Karathanassi, V.,2007. Investigation of the dual-tree complex and shift-invariant discrete wavelet transforms on Quickbird image fusion. *Geoscience and Remote Sensing Letters*, 4(1), pp. 166-170.

Krotkov, E.,1987. Focusing. *International Journal of Computer Vision*, 1, pp. 223–237.

Li, H., Manjunath, B. S., Mitra, S. K.,1995. Multisensor image fusion using the wavelet transform. *Graphical Models and Image Processing*, 57(3), pp. 235-245.

Li, S. T., Ianes, T. K., Wang, Y. N.,2002. Using the discrete wavelet frame transform to merge Landsat TM and SPOT panchromatic images. *Information Fusion*, 3(1), pp. 17-23.

Nayar, S. K., Nakagawa, Y.,1994. Shape from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16 (8), pp. 824–831.

Tyan, J. K.,1997. Analysis and application of autofocusing and three-dimensional shape recovery techniques based on image focus and defocus. PhD. Dissertation, State University of New York, USA.

Wald, L., Ranchin, T., Mangolini, M.,1997. Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images. *Photogrammetric Engineering and Remote Sensing*, 63(3), pp. 691-699.

Wang, H. H., Peng, J. X., Wu, W.,2002. Remote sensing image fusion using wavelet packet transform. *Journal of Image and Graphics*, 7(9), pp. 932-937.

Wang, X. H., Istepanian, R. S. H., Song, Y. H.,2003. Microarray image enhancement by denoising using stationary wavelet transform. *IEEE Transactions on NanoBioscience*, 2(4), pp. 184-189.

Yang, J., Zhao, Z. M.,2007a. Multifocus image fusion based on wavelet packet transform. *Remote Sensing Information*, 1, pp. 11-14.

Yang, J., Zhao, Z. M.,2007b. Multifocus image fusion method based on curvelet transform. *Opto-Electronic Engineering*, 34(6), pp. 67-71.

ACKNOWLEDGEMENTS

This research is funded by China's Special Technical D&R Project for Scientific Academies or Institutes, MOST and Basic Research Fund of Chinese Academy of Meteorological Sciences.