

## Visual and Infrared Image Fusion Based on Contourlet Transform

Chen Zihong<sup>1, a</sup>, Zhang Dexiang<sup>\*1, 2, b</sup>, Yan Qing<sup>1, c</sup>, Zhang Jingjing<sup>1, d</sup>

<sup>1</sup> The School of Electrical Engineering and Automation of Anhui University, Hefei, 230601, China

<sup>2</sup> The School of Electronic and Electrical Engineering of Anhui Sanlian University, Hefei, 230601, China

<sup>a</sup> chenzihong315@126.com, <sup>b</sup> zdxzxy@126.com <sup>c</sup> rubby\_yan5996@sina.com, <sup>d</sup> helenzjj@aiofm.ac.cn

**Keywords:** image fusion; contourlet transform; visual and infrared image; evaluation;

**Abstract.** A new region-based image fusion technique using the contourlet transform (CT) is proposed to produce a fused image better for human and machine interpretation. For good properties of multiscale, localization, directionality and anisotropy, the infrared and visual images are first decomposed using contourlet transform then fusion rules were applied to low frequency components and high frequency components of contourlet coefficients. For the low-frequency coefficients, the average fusion algorithm is used. For the each directional high frequency sub-band coefficients, the region energy and region average gradient maximum criterion are used to select the better coefficients for fusion. Compared with the traditional image fusion algorithms, the fusion algorithm presented in this paper provides better subjective visual effect, and the standard deviation and entropy value would be somewhat increased.

### Introduction

The effort of combining two or more images of the same scene in order to produce a clear image have been proposed to meet the requirements of the different applications in different fields such as concealed weapon detection, remote sensing, and medical imaging. Because of different imaging mechanism and waveband, the infrared image reflects radiation information of the objective scene but image clarity is lower. The visible image reflects reflection information of the objective scene, and can give a better description of surroundings information<sup>[1]</sup>. Through the fusion of the two kinds of the image, combing the complementary information and redundant information of the two kinds of images, the fused image of the target will have clear outline, rich information and perception ability of the image sensor can be largely extended.

The fusion methods of images have been studied widely. Some early methods such as pyramid transform, RoLP pyramid, wavelet transform, and other methods, have gained good fusion effects<sup>[2]</sup>. Wavelet transform, which was developed in the 90s of century, has many advantages in analyzing and processing one dimension-seg-continuous signal<sup>[3]</sup>. However, when it is used in two dimensions(2-D) and higher, 2-D separable wavelet made by one-dimension wavelet only have limited directions which are horizontal, vertical and diagonal. It's not optimal for representing multidimensional signal which include linear or area singularities and has limitations in edge information mining. In 2002, Minh N. Do and Martin Vetterli pioneered a new system of representations named contourlet, which can effectively overcome the disadvantages of wavelet<sup>[4]</sup>.

A new region based image fusion technique is presented in this paper using contourlet transform<sup>[4]</sup>. This transform efficiently deals with images containing contours and texture as is it exhibits high directionality and anisotropy. Two input images are used in this algorithm. After decomposition, select the informative importance measure to dispose the low frequency sub-band and other pass-band directional sub-band. The final fused image is then computed by taking the inverse CT on the composite coefficient representation. A combination of the infrared with a classical visual image of the same scene can be used to detect the target and precisely identify its location<sup>[5]</sup>.

## Contourlet transform

To obtain multi-scale decomposition, the combination of Directional Filter Bank (DFB) and Laplacian Pyramid (LP) is used which is called as Pyramidal Directional Filter Bank(PDFB) [6]. The transform of the CT is achieved via two major steps: The first stage, the Laplace Pyramid (LP) is used to capture the point's discontinuities. The second, the Directional Filter Bank (DFB) is utilized to link the discontinuities into linear structure [7]. Figure 1 shows a multi-scale and directional decomposition using a combination of a LP and a DFB.

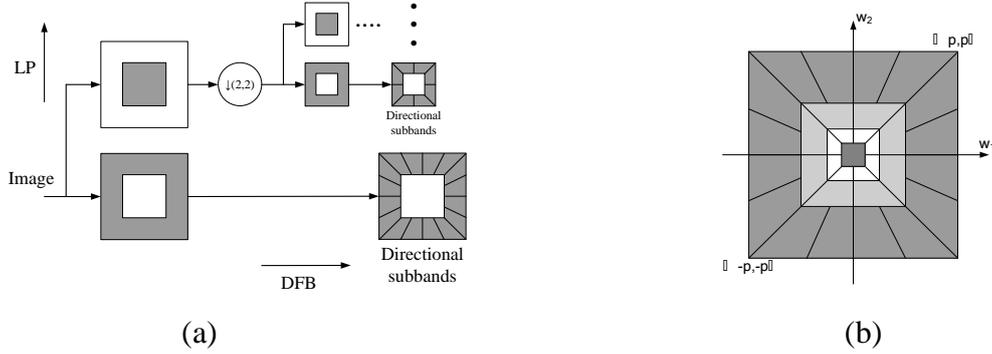


Fig. 1 (a) The frame of contourlet decomposition.(b) Idealized frequency partitioning obtained with the proposed structure.

The main advantage of the CT is the flexibility in choosing the number of directions at each level. Also, the utilization of the iterated-filter banks increases its computational efficiency. Contourlet transform can provide detailed information in any arbitrary direction, as the number of directional subbands in each scale is usually  $2^n$  ( $n \in \mathbb{N}$ ), which is quite flexible when different  $n$  is chosen.

## Visual and infrared image fusion algorithm

In this section, an adaptive visual and infrared image fusion method is presented based on contourlet transform. Firstly the source images which have been perfectly co-aligned are decomposed at different scales and directions by contourlet transform. Thus we can obtain the low frequency sub-band coefficients  $C_{j_0}^A(x, y)$  and  $C_{j_0}^B(x, y)$  and various band-pass directional sub-band coefficients  $C_{j,l}^A(x, y)$  and  $C_{j,l}^B(x, y)$ . Considering the characteristics of decomposition subbands, adopt different fusion rules to low-frequency subbands and high-frequency directional subbands. Finally through reconstructing the coefficients by inverse CT, we get the fused image  $F$  [8].

The low-frequency sub-band reflects the general information of the image. So we compute the average value of the two sub-band coefficient as the low-frequency sub-band coefficient of the fused image.

$$C_{j_0}^F(i, j) = (C_{j_0}^A(i, j) + C_{j_0}^B(i, j)) / 2 \quad (1)$$

Where  $C_{j_0}^A(i, j)$  and  $C_{j_0}^B(i, j)$  respectively represent the low-pass decomposition coefficients of image A and B, and  $C_{j_0}^F(i, j)$  is the coefficients of fused image F.

The high (band) frequency includes most of the image detail information (edge and texture information). As for the high frequency sub-band's coefficient, in the largest decomposition level  $j$ , The fusion rules compute the weighted region energy of the center pixel and its neighboring region energy of the center pixel. Define the region energy  $E(i, j)$  centered on the current coefficient [9].

$$E(i, j) = \sum_{m=-ln}^1 \sum_{n=-1}^1 w(m, n) [I(i+m, j+n)]^2 \quad (2)$$

Where  $w(m,n)$  define a  $3 \times 3$  window containing neighboring coefficient, and  $I(i,j)$  presents the value of the pixel. So in this new fusion scheme, weighted region energy fusion rule is adopted for band-pass coefficient.

$$C_{j,l}^F(i,j) = \begin{cases} C_{j,l}^A(i,j) & E_{j,l}^A(i,j) \geq E_{j,l}^B(i,j) \\ C_{j,l}^B(i,j) & E_{j,l}^A(i,j) \leq E_{j,l}^B(i,j) \end{cases} \quad (3)$$

As for other  $j-1$  lever high frequency coefficient, the maximum principle of the local average gradient has been adopted; the local average gradient is defined as follows <sup>[10]</sup>.

$$G_{j,l}^l(m,n) = \frac{1}{MN} \sum_{m=-\frac{M-1}{2}}^{\frac{M-1}{2}} \sum_{n=-\frac{N-1}{2}}^{\frac{N-1}{2}} \sqrt{\frac{\Delta S_x^2(m,n) + \Delta S_y^2(m,n)}{2}} \quad (4)$$

Where,  $\Delta S_x$  and  $\Delta S_y$  are respectively the variance of horizontal and vertical directions. That is to say, in the transform domain, calculating the corresponding point  $M \times N$  neighborhood's average gradient, and choosing the higher coefficient of the variance as the fused image's corresponding point's coefficient.

$$C_{j,l}^F(x,y) = \begin{cases} C_{j,l}^A(x,y) & |G_{j,l}^A(x,y)| \geq |G_{j,l}^B(x,y)| \\ C_{j,l}^B(x,y) & |G_{j,l}^A(x,y)| < |G_{j,l}^B(x,y)| \end{cases} \quad (5)$$

## Experiments and Results

The paper has adopted several different approaches to fuse the two strictly matched infrared and visible images. The figure 2 presents a group of image fusion results which is based on multi-image fusion approach. Figure 2 (a) and (b) are the source images. The figure 2(c) to figure 2 (e) are respectively the fused images based on the Contrast Pyramid decomposition, Laplacian Pyramid decomposition and Discrete wavelet transform. Figure 2 (f) is a clear image with rich features using CT method. Among these, the approach LP, DWT and CT have adopted the same fusion rule as this paper. In calculating the local variance, the  $M \times N = 3 \times 3$ . The decomposition of LP, DWT and CT are all 3 levels.

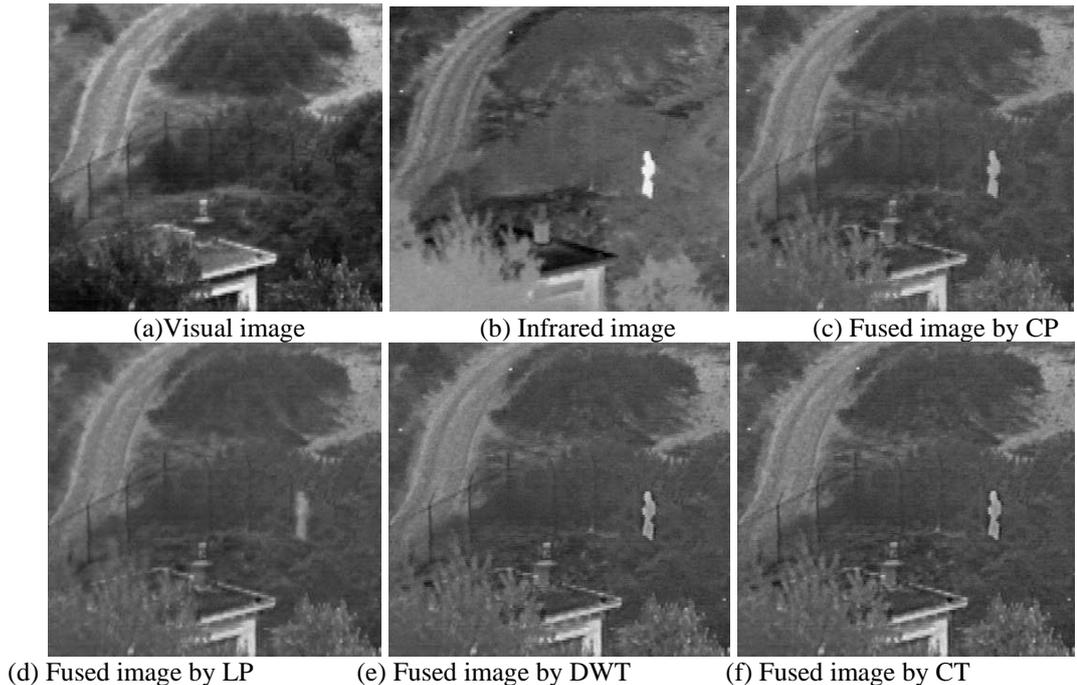


Figure 2: Comparison of image fusion based on different rules

In order to objectively assess the quality of the fused image, the paper adopted the evaluation criteria of average value, entropy, standard deviation, the mean cross entropy and the root cross entropy, which not only measure distortion in the structure of two images, but also can evaluate the retained significant information such as variance, edge and contrast. The comparison is shown in Table1.

Table 1 Criteria value of fused images by different methods

Criteria method	average value	standard deviation	entropy	the mean cross entropy	the root cross entropy
CP	103.578	28.450	6.608	-0.831	1.078
LP	103.576	28.106	6.671	-0.876	1.086
DWT	103.582	28.532	6.682	-0.818	0.970
CT	103.578	28.537	6.690	-0.801	0.921

From Table 1, compare with other transform, shows the statistical parameters of different methods. The fused result with DWT held higher average value. But the result with CT held higher value of entropy and the MCE and RCE are the smallest in the four methods. It suggest that the CT fusion image can improve spatial resolution and is the best in the given methods.

## Conclusion

In this paper, a novel image fusion method based on contourlet transform is presented. The algorithm can fuse corresponding information in different resolutions and directions, In order to obtain clearer contour and more details, spatial frequency is used to select adapt coefficient of low-pass and high-pass subbands in CT domain. Compared with the fusion method based on wavelet and other methods, the experiment shows that this approach can achieve better result than them.

## Acknowledgment

This work was financially supported by the Chinese National Science Foundation Grant (No.61272025) and the Youth Scientific Research Foundation of Anhui University 211 Project (KJQN1114).

## References

- [1] Weihua he, Yongcai Guo, ChaoGao, and Dongguo Zhou. International Conference on computer assist design and image. Vol. 5 (2011), pp.884.
- [2] Burt , P.J.and Adelson, E.H. EEE Transaction on communication. Vol. 31 (2003), pp. 532.
- [3] Urbach, E.R, Roerdink, J.B.T.M, Wilkinson, and M.H.F, Pattern Analysis and Machine Intelligence. Vo.29 (2007), pp. 272.
- [4] Minh.N.Do, Dept.of Elet. and Comp.Eng, Systems and Computer. Vo.1 (2002), pp. 497.
- [5] Cunha A L,Zhou J P, Do M N, IEEE Transactions on Image Processing. Vol. 15 (2006), pp. 3089.
- [6] D.D.-Y. Po and M.N. Do, IEEE Transactions on Image Processing. Vol. 15 (2006), pp. 1610.
- [7] J.Zhou, A L.cunha, and M N. Do, IEEE International Conference on Image Processing. Vol. 12(2005), pp. 469.
- [8] Hai Wang, Zijun Feng, ICFCC Future Computer and Communication. Vol. 10 (2010), pp. 220.
- [9] Wenzhan Dai, Libo Tan, Aiping Yang, Intelligent Control and Automation. Vol.10(2012), pp. 4576
- [10] Lewis, J.J, Nikolov, S.G, Canagarajah, C.N, Bull, D.R, and Toet.A, IEEE International Conference on Image Processing. Vol. 12(2006), pp. 8.