

Image Fusion Algorithm Based on Deng Correlation and Image Saliency in Curvelet Framework

LIU YUANLI^{1, a}, YANGYI^{1, b, *} and SUNJIAN^{1, c}

¹SKLSVMS, School of Aerospace, Xi'an Jiaotong University, Xi'an, Shaanxi, China 710049

^alyl9126@stu.xjtu.edu.cn, ^bjiafeiyy@mail.xjtu.edu.cn, ^csunjian10@mail.xjtu.edu.cn

Keywords: image fusion, saliency, curvelet transform, Deng correlation

Abstract. In this article, a multi-sensor image fusion algorithm is presented. The fusion framework is built on the second discrete curvelet transform to make full use of its multi-scale and multi-direction representability. The fusion rules are mainly based on Deng correlation for representing image similarity and the image saliency for showing the importance of targets. Subjective and objective evaluation results from experimental simulation show the efficacy of the proposed fusion algorithm.

Introduction

Image fusion is a technique utilizing two or more images of the same scene from various sensors to realize the information complementarity and to obtain precise, comprehensive and reliable description of the scene. Image fusion methods can be categorized into three levels, i.e., pixel level, feature level and decision level[1]. Recently, image fusion has been applied to many fields, such as remote sensing, military and medical science, etc. And it can provide exact and clear information for geological analysis, military defence and disease diagnosis. Infrared image and visual image fusion, laser image and visual image fusion and multi-focus images fusion are the three typical image fusion types. The infrared image is sensitive to temperature and can always obtain clear objects even in bad illumination or being sheltered[1]. But it has low contrast, poor visual effect and is lack of spectrum information. The visual image has rich spectrum information, high resolution and wide dynamic range. It is also consistent with human eyes' interpretation[1]. But the visual image has poor quality due to the lack of brightness or the terrible weather. The laser image has high resolution, high signal to noise ratio and distance information, but its color is single and details are blurry[2]. The multi-focus images have different clear regions due to different foci[3]. Owing to information complementarity of different kinds of images, thus we utilize them synthetically to improve image interpretation ability by fusion process.

The inchoate image fusion algorithms include the principal component analysis(PCA) algorithm[1] and the intensity-hue-saturation(IHS) algorithm[1], etc. With the arising of multi-resolution theory, the pyramid decomposition[1], wavelet transform[1] and curvelet transform[1] come into being. These methods can correspond to the human visual system because of their ability in finding the image scales and spatial features simultaneously. In this paper, we use the second discrete curvelet transform for designing fusion framework because the curvelet transform's multi-scale and multi-direction representability can deal with the edge and texture details better.

Besides the fusion framework, the fusion rules are also important. The traditional rules such as the weighted-based rule and the correlation-based rule have poor performance in image contrast and even lead to missing edges. Thus we introduce Deng correlation for dealing with neighborhood information specially and image's global saliency for representing the regions catching first by human eyes into making rules. Thus, we can take the region information and the global saliency information into consideration. The experiments verify the validity of our proposed method.

The Second Discrete Curvelet Transform. Curvelet transform was first proposed by Candes[3]. The essence of curvelet transform is a block multi-scale analysis of the sub-bands that combines the

features of isotropic and multi-scale. The second discrete curvelet transform is based on the second continuous curvelet transform. It is a milestone for the two-dimensional signal analysis and is used to many fields, such as image denoising, image compression, image enhancement, image fusion, image recognition, etc[3].

The definition of the second discrete curvelet transform[3] is as follows:

$$c(j, l, k) = \int f(w)U_j(S_\theta^{-1}w)e^{i(S_\theta^{-1}b, w)}dw = \int f(S_\theta w)U_j(w)e^{i(b, w)}dw \quad (1)$$

$$U_j(w) = W_j(w)V_j(w) \quad (2)$$

$$S_\theta = \begin{pmatrix} 1 & 0 \\ -\tan \theta & 1 \end{pmatrix} \quad (3)$$

Here, W_j is diameter widow function. V_j is angle widow function. j is scale of the decomposed image. S_θ is shear matrix, b belongs to $(k_1 * 2^{-j}, k_2 * 2^{-j/2})$. $\theta_l = 2\pi \cdot 2^{-\lfloor j/2 \rfloor}$. l is uniform rotation angle sequence. $l = 0, 1, \dots$, $k = (k_1, k_2) \in Z^2$ is the translation parameter sequence. W is the frequency variation.

The traditional features of wavelet transform can only provide the features of points, but by definition, the second discrete curvelet transform can provide sparse representation for both edges and textures of the image. And due to the highly directional identification and anisotropy of the second discrete curvelet transform, it can describe the features directly of the line and curve, and even the plane[4]. There are two ways to realize the second discrete curvelet transform: unequally spaced FFT (usFFT) and wrapping-based transform (wrap)[4].

Deng Correlation. Deng correlation was first proposed by Julong Deng. It belongs to gray correlation system[5]. Its essence is to compare the curve shapes of the reference sequence and the source sequences. The sequences' similarity can be described by the gray correlation. Recent years, Deng correlation is used to image processing to compute the images' similarity and analyse the images' characteristics. It can be established by the following five steps[5]:

First, we introduce the the reference sequence y_0 and the source sequences y_i ($i=1, \dots, n$). Second, if the sequence's dimension is different, we also need the data dimensionless. Third, we need to find the maximum and minimum data of the sequences that have been subtracted. Fourth, we compute the Deng correlation by Eq.4. Fifth, compute the correlation coefficient R_d by Eq.5.

$$\xi_{i,0}(k) = \frac{\min_i \min_k |y_0(k) - y_i(k)| + \rho \max_i \max_k |y_0(k) - y_i(k)|}{|y_0(k) - y_i(k)| + \rho \max_i \max_k |y_0(k) - y_i(k)|} \quad (4)$$

$$R_d(y_i, y_0) = \frac{1}{n} \sum_{i=1}^n \xi_{i,0}(k) \quad (5)$$

Here, ρ is the weighted coefficient. $\rho \in (0,1)$, $y_0(k)$ is the k -th data of y_0 . $y_i(k)$ is the k -th data of y_i . Because Deng correlation is the most classical method to compute the gray correlation, and it has superiority for dealing with the small regional neighborhood information, it can be introduced to analyse the neighborhood pixels sequences of images. We combine it with curvelet transform and then we use it to make the fusion rule for the low frequency sub-images.

Saliency Map. In this article, the Itti's model[6] is used to compute saliency maps and then to compute the weights for fusion. Because it is capable of strong performance for complex natural scenes, despite that its architecture and feed-forward feature-extraction mechanisms are simple. The saliency map can be obtained by three steps: features extraction, feature maps normalization and feature maps fusion.

First, an intensity image I is defined as $I = (R + G + B)/3$, with R , G , and B being the red, green, and blue channels of the input image. I is used to create a Gaussian pyramid $I(p)$, where p belongs to the

range of [0..4] and p is the scales. The R , G , and B channels are normalized by I in order to decouple hue from intensity[6]. Then we use the center-surround differences mechanism between a “center” fine scale c and a “surround” coarser scale s to yield the feature maps. Here, c is 1 scale and s is belonging to $\{2,3\}$. Finally, we compute the intensity map by the Eq.6. Second, we utilize Gabor pyramids $O(p,q)$ to obtain the local orientation information, where p belongs to the range of (0..4), and q is the angles that belongs to $\{0,45,90,135\}$. Then the orientation maps are computed by the Eq.7. Third, they are normalized by dividing the maximum of the feature maps, after that they are summed up for getting intensity I_{norm} and orientation O_{norm} feature maps. Finally, we get the saliency map S by the Eq.8.

$$I(c, s) = |I(c) - I(s)| \quad (6)$$

$$O(c, s, q) = |O(c, q) - O(s, q)| \quad (7)$$

$$S = (I_{norm} + O_{norm}) / 2 \quad (8)$$

Here, $||$ is the absolute function.

In this paper, we compute the original images’ saliency maps and use them to compute the fusion weights. Thus, we can utilize the global information to make sure the complementary information being retained comprehensively.

The Proposed Fusion Algorithm. Due to the second discrete curvelet transform’s advantages in dealing with the edges and textures, our fusion framework is based on it to decompose the original images to n scales and then each scales’ coefficient matrixs can be obtained. Fusion rule is made by Deng correlation of low frequency sub-images and regional energy maximum of high frequency sub-images. The fusion scheme is shown in Fig.1. The fusion steps can be described as the following:

First, we utilize the curvelet transform to decompose the registered original image1,2 into n levels, then the low and high frequency sub-images can be obtained. Second, we construct low frequency sub-images fusion rules. Low frequency sub-images mainly contain color, intensity and brightness information. We make the rule by computing the low frequency sub-images’ Deng correlation R_d . If $R_d > th$ (threshold, $th \in (0,1)$), we use Eq.6-Eq.8 to compute the saliency maps (S_1 and S_2), after that we compute the weight w by Eq.9. Otherwise, we use the R_d as the weight by Eq.10. Third, we construct high frequency sub-images fusion rule. High frequency sub-images contain the edge and texture information. In order to remain the details as many as possible, we select the high frequency coefficient matrices by comparing their regional energy, then we select the lager coefficient matrix by Eq.11. Finally, we utilize the curvelet inverse transform to get the final fusion result.

$$w = \frac{mean(S_1)}{mean(S_1) + mean(S_2)} \quad (9)$$

$$\begin{cases} R_d > th, F_l = w * I_{1l} + (1 - w) * I_{2l} \\ R_d \leq th, F_l = I_{1l} * (1 - R_d) + I_{2l} * R_d \end{cases} \quad (10)$$

$$\begin{cases} F_h = I_{1h}, \text{if} (energy(I_{1h}) \geq energy(I_{2h})) \\ F_h = I_{2h}, \text{if} (energy(I_{1h}) < energy(I_{2h})) \end{cases} \quad (11)$$

Here, F_l , I_{1l} , I_{2l} respectively are low frequency coefficients for the fusion result, image1(I_1) and image2(I_2). l is low frequency. F is the fusion image. F_h , I_{1h} , I_{2h} respectively are high frequency coefficients for the fusion result, image1(I_1) and image2(I_2). h is high frequency.

Experimental Results. In order to verify the efficacy of the proposed fusion method, three sets of registered multi-sensor images are used for experiments. They are infrared image and visual image,

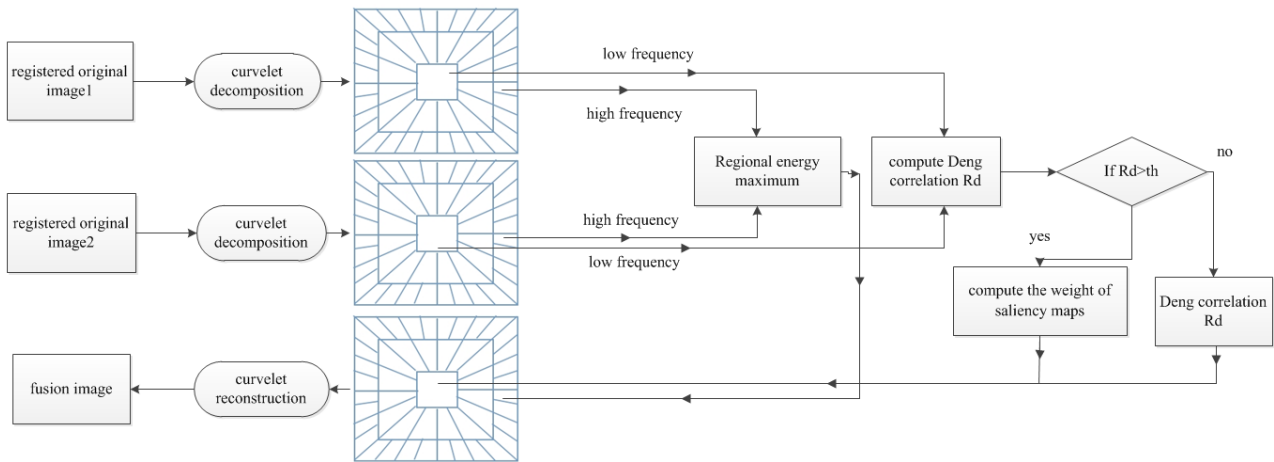


Fig.1. The flowing chart of the proposed fusion algorithm

laser image and visual image and left-focus image and right-focus image. We compare the proposed fusion method (PF, curvelet decomposition level is 6, decomposition method is wrap) with principal component analysis algorithm (PCA), gradient pyramid algorithm(GP, decomposition level is 3), wavelet transform algorithm (WT, wavelet basis is sym4 and decomposition level is 3) and curvelet transform algorithm (CT, decomposition method is wrap, decomposition level is 6). Experimental results are presented in Fig.2, Fig.3 and Fig.4. Fig.2 is the fusion results for infrared image and visual image (image size both are 512*512, $th=0.7$). Fig.3 is the fusion results for laser image and visual image (image size both are 256*256, $th=0.5$). Fig.4 is the fusion results for multi-focus images (image size both are 512*512, $th=0.9$).

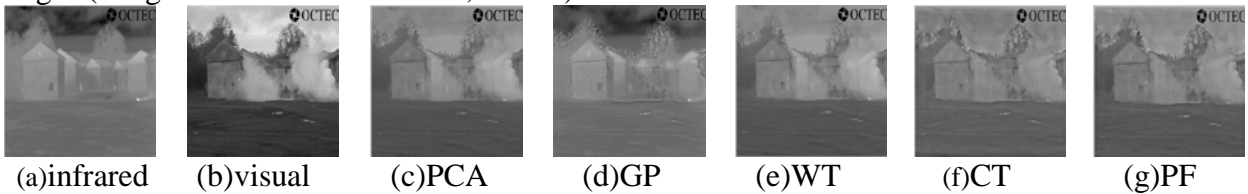


Fig.2. Fusion result of infrared image and visual image

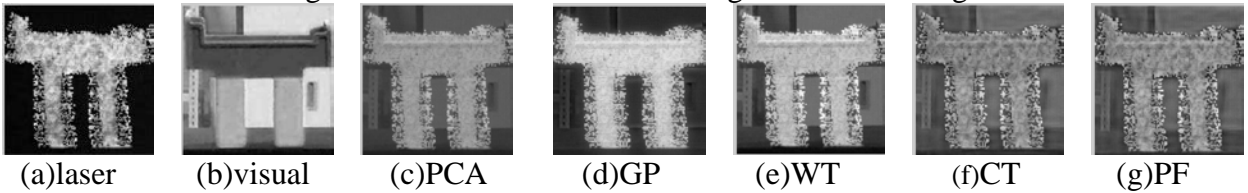


Fig.3. Fusion result of laser image and visual image

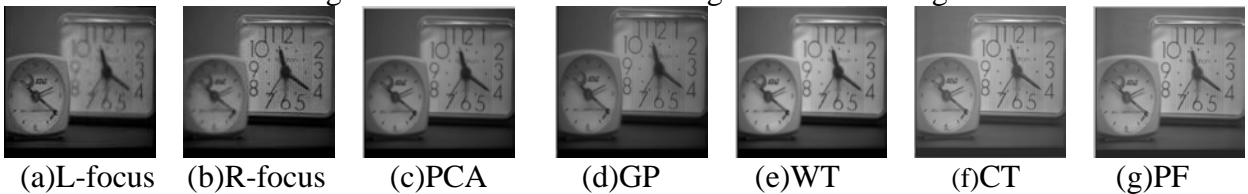


Fig.4. Fusion result of multi-focus images

From the visual effect, the fusion images of CT and PF are better than the PCA, GP and WT in retaining of edges, textures and color information. The luminance of the fusion image's background is worst in GP, and the effect of GP is the most unstable. PCA and WT are two methods whose performance are better than GP in background's brightness and color information retaining. The performance of CT and PF are mostly identical with minor differences. In addition, we compare the five algorithms by three objective indexes: average gradient[1] (AG), spatial frequency[1] (SF) and weighted edge information retaining[1] (Q). The results are shown in Table1.

If AG is larger, the fusion images will be more legible. If SF is larger, the fusion images will be better. If Q is larger, the more edges information will be retained and the fusion image will be much better. From Table1, we can conclude that the proposed fusion algorithm has a certain improvement whatever AG, SF and Q. This conclusion is in agreement with the visual effect. Thus the experiments confirm the effectiveness of the proposed algorithm.

Table1. The objective indexes of different algorithms

three set of images	algorithms	objective indexes		
		AG	SF	Q
infrared image visual image	PCA	0.0136	9.7779	0.4164
	GP	0.0205	14.8157	0.5741
	WT	0.0135	9.7063	0.3987
	CT	0.0233	16.8313	0.6016
	PF	0.0233	16.8316	0.6116
laser image visual image	PCA	0.0596	21.4153	0.5255
	GP	0.0966	34.7131	0.6727
	WT	0.0541	19.4486	0.4158
	CT	0.1062	38.1627	0.6714
	PF	0.1063	38.1754	0.6747
left-focus image right-focus image	PCA	0.0071	5.0879	0.5863
	GP	0.0090	6.4830	0.6411
	WT	0.0071	5.0864	0.5849
	CT	0.0112	8.1060	0.6441
	PF	0.0112	8.1060	0.6641

Summary

This article utilize the advantage of the second discrete curvelet transform which can deal with the edge and texture information better and has good directionality. In the curvelet framework, the Deng correlation and image saliency are combined for a new fusion rule construction dealing with images similarity and global saliency information. Visual effect and objective indexes both verify the proposed method is effective.

Acknowledgements

This work was supported by the National Natural Science Foundation (No. 61573275, No. 61203222), Science and technology project of Shaanxi Province (No. 2013KJXX-46), Specialized Research Fund for the Doctoral Program of Higher Education (20120201120036), and Fundamental Research Funds for the Central Universities (xjj2014122).

Reference

- [1]Xi Cai: *Multi-Scale Image Fusion Theory and Method* (Publishing House of Electronics Industry, Beijing 2014)
- [2]S.Frintrop, E.Rome, A.Nuchter and H.surmann: A Bimodal Laser-Based Attention System. *Computer Vision and Image Understanding* Vol.100 (2005), p.124-151
- [3]M.Miao, W.Renyuan: Multi-focus Image Fusion Based on Similar Gray Correlation in Curvelet Domain. *Electronic Journals* Vol.40 (2012), p.1984-1988
- [4]J.Xiaotao: *Infrared Image and Visual Image Fusion Based on Curvelet Transform*. (Xi Dian University, 2012)
- [5]D.Julong: *The Basic Method of Gray System*. (Publishing House of Wuhan University, 2005)
- [6]L.Itti, K.Christof and N.Ernst: A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence* Vol.11 (1998), p.1254-1259