Journal of Zhejiang University SCIENCE A
ISSN 1009-3095
http://www.zju.edu.cn/jzus
E-mail: jzus@zju.edu.cn

Multiresolution image fusion scheme based on fuzzy region feature*

LIU Gang (刘 刚)[†1,2], JING Zhong-liang (敬忠良)[†1], SUN Shao-yuan (孙韶媛)[1]

([1]School of Electrical and Information Engineering, Shanghai Jiao Tong University, Shanghai 200030, China)

([2]School of Electric Power and Automation Engineering, Shanghai University of Electric Power, Shanghai 200090, China)

[†]E-mail: lukelg@sjtu.edu.cn; zljing@mail.sjtu.edu.cn

Received July 10, 2005; revision accepted Dec. 4, 2005

Abstract: This paper proposes a novel region based image fusion scheme based on multiresolution analysis. The low frequency band of the image multiresolution representation is segmented into important regions, sub-important regions and background regions. Each feature of the regions is used to determine the region's degree of membership in the multiresolution representation, and then to achieve multiresolution representation of the fusion result. The final image fusion result can be obtained by using the inverse multiresolution transform. Experiments showed that the proposed image fusion method can have better performance than existing image fusion methods.

Key words: Image fusion, Image multiscale decomposition, Discrete wavelet frame

doi:10.1631/jzus.2006.A0117 **Document code:** A **CLC number:** TP391.4

INTRODUCTION

Image fusion can be defined as the process by which several images, or some of their features, are combined together to form different modalities or instruments, and is of great importance in many applications (Hall and Llinas, 1997; Zhou *et al*., 1998), such as object detection, ATR (Automatic Target Recognition), remote sensing, computer vision, and robotics.

Recently some researchers recognized that it seems more meaningful to combine objects/regions rather than pixels. Zhang and Blum (1997) proposed a region-based fusion algorithm which combines images guided by the identification of important features (such as object and regions of interest) in each region. Piella (2002) also proposed a region-based fusion scheme. Because the region feature is obtained at each level of the multiresolution representation, although their region feature has some improvements, the selection result of the high frequency bands is different in that the final fusion result cannot keep the region feature.

Therefore, this paper proposes an algorithm based on fuzzy region feature and implementing the fusion process in fuzzy space. In this paper, we just focus on the pixel gray level distribution in the image region in order to preserve the region feature of the source images. The rest of this paper is organized as follows. Section 2 presents a new framework for image fusion. The main contributions to image fusion can be found in this section. Section 3 describes the discrete wavelet frame transform. Experimental results and their comparison with those in the literature are reported in Section 3. Section 4 concludes the paper.

*Project supported by the National Natural Science Foundation of China (No. 60375008), China Aviation Science Foundation (No. 02D57003), China Ph.D Discipline Special Foundation (No. 20020248029), and Shanghai Key Scientific Project (No. 02DZ15001), China

NEW IMAGE FUSION ALGORITHM

The paper adopts the multiresolution analysis discrete wavelet frame transform and fuzzy region feature fusion scheme to implement the selection of source image wavelet coefficients. Fig.1 is the framework of the proposed image fusion algorithm. The first step is to choose an image as object image that can reflect the object and background clearer than the other image. The second step is to decompose the source image into multiresolution representation. There are low frequency band at each level during the next level decomposition. The low frequency bands of the object image are segmented into region images. The third step is defining the attributes of the regions by some region features, such as the mean of gray level in a region. In this case, each pixel point has its membership value. Then using certain attribute region fusion scheme combining with the membership value of each pixel, the multiresolution representation of the fusion result is achieved using defuzzification process. The final step is to do inverse discrete wavelet frame transform, and the final fusion result is obtained.

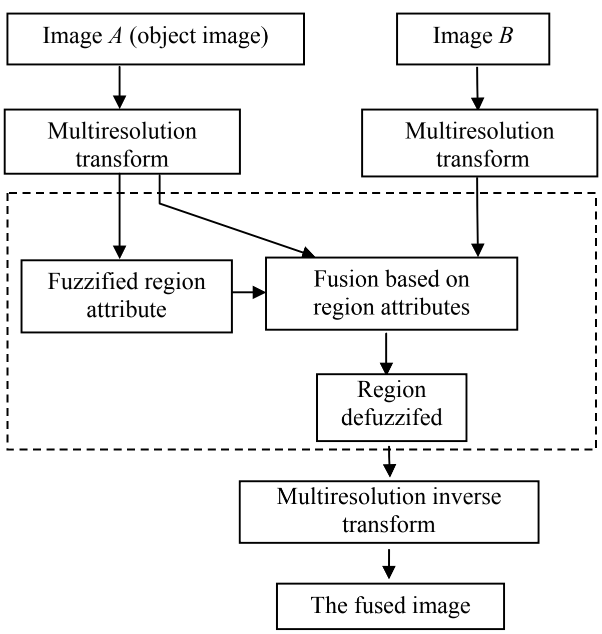

Fig.1 Framework of the proposed image fusion algorithm

Discrete wavelet frame

As we know, the discrete wavelet transform is shift dependent. This property may lead to some defects in image fusion. To overcome the shift dependency of the wavelet fusion method, the input images must be decomposed into a shift invariant wavelet representation. There are several ways to achieve this. An approach, related to the concept of wavelet frames, was proposed by Unser (1995).

As in the DWT, each stage of the DWF splits the input sequence into the wavelet sequence $w_t(n)$ and the scale sequence $I_t(n)$ to be used as input for the next decomposition level. That is,

$$w_{t+1}(n) = \sum_k g(2^t k) I_t(n-k), \tag{1}$$

$$I_{t+1}(n) = \sum_k h(2^t k) I_t(n-k), \tag{2}$$

where $g(\cdot)$ and $h(\cdot)$ are the high and low frequency analysis filters respectively, n denotes the pixel position, k represents the shift of the position, and t denotes the scale of the DWF. The zeroth level scale sequence is set equal to the input sequence $f(n)$, i.e., $I_0(n)=f(n)$ is input signal. In contrast to the standard DWT, the subsampling is dropped which results in a highly redundant wavelet representation. The analysis filters $g(2^t k)$ and $h(2^t k)$ at level t are obtained by inserting an appropriate number of zeros between the filter taps of the prototype filters $g(k)$ and $h(k)$.

The input sequence is reconstructed by the inverse DWF recursively as a convolution of the wavelet and scale sequences with the appropriate reconstruction filters $\tilde{g}(2^t k)$ and $\tilde{h}(2^t k)$:

$$I_t(n) = \sum_k \tilde{h}(2^t n - k) I_{t+1}(n) + \sum_k \tilde{g}(2^t n - k) w_{t+1}(n). \tag{3}$$

The reconstruction filters $\tilde{g}(k)$ and $\tilde{h}(k)$ can be achieved by using the relation of conjugate quadrature filter. That is,

$$H(z)\tilde{H}(z^{-1}) + G(z)\tilde{G}(z^{-1}) = 1, \tag{4}$$

where H, $\tilde{H}$, G, and $\tilde{G}$ are the frequency domain counterparts of the filters h, $\tilde{h}$, g, and $\tilde{g}$ respectively.

Due to the averaging inherent in both the discrete wavelet frame transform and the inverse discrete wavelet frame transform, the filter coefficients from the standard DWT scheme must be slightly modified to obtain perfect reconstruction, i.e. energy normalization has to applied to the filters:

$$h^{\mathrm{DWF}}(k)=\frac{1}{\sqrt{2}}h^{\mathrm{DWT}}(k), \tag{5}$$

$$h^{\mathrm{DWF}}(2^t k)=\frac{1}{\sqrt{2}}h^{\mathrm{DWF}}(2^{t-1}k), \tag{6}$$

When applied this modification, all other results concerning the filters, i.e. quadrature mirror condition, biorthogonality condition and the relation between low-pass and high-pass filters can be adopted from the DWT case.

Fusion of certain attribute region

In this paper, the source images are segmented into important regions, subimportant regions and background regions. We choose an object image from the source images to perform segmentation and to define region attribute. The segmentation procedure is using *K*-Means clustering algorithm according to the pixel gray level distribution. For the detail of the algorithm are given by Weisstein E.W. (http://mathworld.wolfram.com/K-MeansClusteringAlgorithm.html). *K*-Means clustering algorithm generates a specific number of disjoint, flat (non-hierarchical) clusters. The object image is segmented into important regions, subimportant regions and background regions. The segment result is shown in Fig.2. The white portion represents important region, the black one represents the background region and the gray one represents the subimportant region.

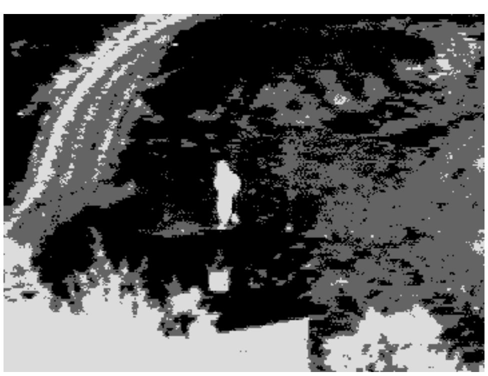

Fig.2 The segmentation result of the IR image in Fig.6a using *K*-Means clustering algorithm

If the region is important to the background, the region feature is important. So the region should be selected as fusion result as directly as possible. If the region feature indicates the region is important, we select the region of image *A*. If the region feature indicates the region is background, we select the region of image *B*. If the region is subimportant, the region importance is equivalent to that of the two source images; the single pixel salient feature is more important than region feature (Zhang and Blum, 1997). It is not very important to preserve the region feature, so the more salient pixels are selected in multiresolution space as the fusion result. According to the above analysis, three fusion schemes are involved at the same time. They correspond to the three kinds of image regions.

Since the importance of image region is relative, that is to say, we cannot very definitely determine a region is important or not. The region feature importance is a fuzzy concept. Therefore, it is necessary to fuzzy the importance attribute of regions. The fusion process is performed in fuzzy space.

Function of membership and defuzzification

Suppose the highest gray level and lowest gray level is $L_{\max}$ and $L_{\min}$ respectively, default is 255 and 0. Then defining the function of membership *i* region belongs to *j* is:

$$\mu_{i,j}=\exp\left[\frac{-(ME_i-E_j)^2}{(L_{\max}-L_{\min})/2}\right], \tag{7}$$

where $E_1=L_{\min}$, $E_2=(L_{\max}-L_{\min})/2$, $E_3=L_{\max}$; and ME_i is the mean of pixel gray level within region *i*; $\mu_{i,1}$, $\mu_{i,2}$, $\mu_{i,3}$ are the values of membership in important region, subimportant region and background region.

In the sketch map Fig.3, E_1, E_2 and E_3 are the three respective attributes of the image important region, subimportant region and background region. $ME_i=E_1$ indicates that the region *i* is background, the fusion result F_1 is the corresponding region of image *B*; $ME_i=E_2$ indicates that the region *i* is subimportant, the fusion result F_2 is obtained by single pixel based fusion algorithm; $ME_i=E_3$ indicates that the region *i* is important, the fusion result F_3 is the corresponding region of image *A*.

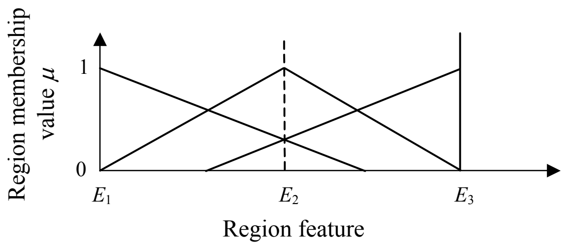

Fig.3 Sketch map of the curve of region membership value

Finally, depending on the feature of every region, the membership of each pixels is defined as $\mu_{i,1}$, $\mu_{i,2}$, $\mu_{i,3}$. The final fusion result is achieved by defuzzification process using the membership:

$$F=\sum_{i=1}^{3}\mu_{i,j}F_j\Big/\sum_{i=1}^{3}\mu_i\ , \tag{8}$$

where F is the multiresolution representation of the fused image of Figs.6a and 6b (Fig.4). The final fused image is obtained by performing inverse discrete wavelet frame transform.

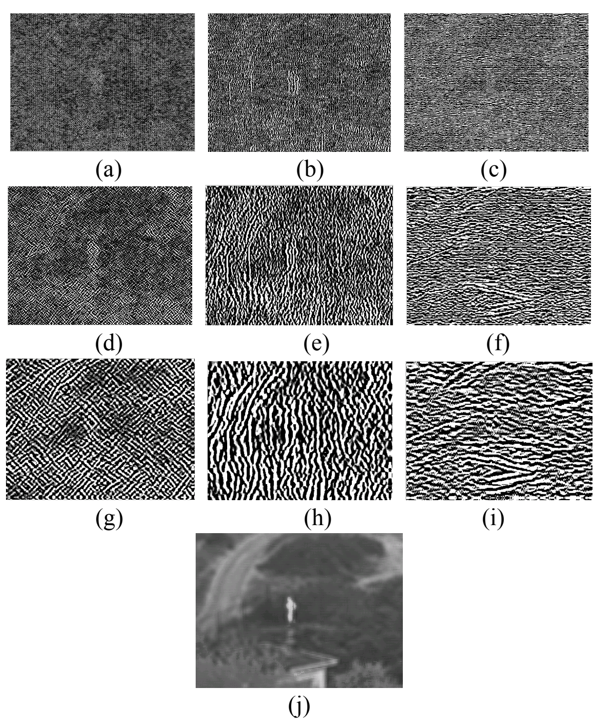

Fig.4 The multiresolution representation of the fused image of Figs.6a and 6b. (a)~(i) the high frequency bands, (j) the lowest frequency band

EXPERIMENT AND DISCUSSION

In this paper, two evaluation criteria are used for quantitatively assessing the performance of the fusion. The first evaluation measure is the objective performance metric that is proposed by Xydeas and Petrovic (2000). It models the accuracy with which visual information is transferred from the source images to the fused image. Therefore, we call it edge mutual information.

Mutual information has been proposed for fusion evaluation. Given two images x_F and x_R we define their mutual information as

$$I(x_R;x_F)=\sum_{u=1}^{L}\sum_{v=1}^{L}h_{R,F}(u,v)\log_2\frac{h_{R,F}(u,v)}{h_R(u)h_F(v)}\ , \tag{9}$$

where x_R is the ideal reference, x_F is the obtained fused image, h_R, h_F are the normalized gray level histograms of x_R, x_F respectively, $h_{R,F}$ is the joint gray level histogram of x_R and x_F, and L is the number of bins. We select L=100. Thus, the higher the mutual information between x_R and x_F, the more likely x_F resembles the ideal x_R. Mutual information evaluation method may be modified into an objective measure according to (Qu *et al*., 2002). This is the second evaluation measure.

CCD image and MMW image fusion, CCD image and infrared image and medical image CT and MRI image were studied in the work described in this paper.

Figs.5a and 5b are CCD images and an MMW image employed in Concealed Weapon Detection. In this kind of images, the region feature is defined by the MMW image. We can find that a weapon is concealed in the third person from Fig.5b. Zhang Zhong's wavelet based algorithm (Zhang and Blum, 1997), the Laplacian based algorithm of Piella (2002), discrete wavelet frame based algorithm (Rockinger and Fechner, 1998) and the algorithm proposed in this paper were separately implemented on the fusion process. The decomposition level was chosen as 3. The fusion result is shown as Figs.5c~5f. It is clear that the proposed method outperformed the others. The quantity of the evaluation measure is shown in Table 1, where the numbers of pixel mutual information and edge mutual information go up.

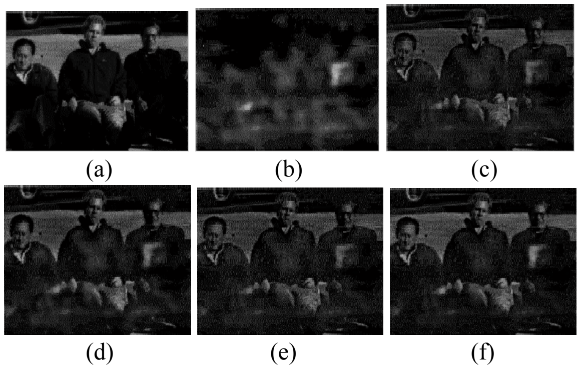

Fig.5 Fusion result of visual image and MMW image. (a) CCD image; (b) MMW image; (c) Zhang Zhong's algorithm; (d) G. Piella's algorithm; (e) Discrete wavelet frame; (f) Algorithm proposed in this paper

Table 1 Fuse result of CCD image and MMW image

Evaluation	PMI	EMI
DW	0.2429	0.4993
LP	0.2800	0.6011
DWF	0.2525	0.6036
Proposed algorithm	0.3422	0.6169

Fig.6b shows a visual image of a scene with the background consisting of road, grass land, fence and house. But no person appears on the infrared image in Fig.6a. In this kind of images, the region feature is defined by the IR image. Zhang's wavelet based algorithm (Zhang and Blum, 1997), Piella's Laplacian based algorithm (Piella, 2002), discrete wavelet frame based algorithm (Rockinger and Fechner, 1998) and the algorithm this paper proposed were separately implemented on the fusion process. The decomposition level was chosen as 3. The fusion result is shown in Figs.6c~6f. It is clear that the proposed method outperformed the others. The quantity of the evaluation measure is shown in Table 2, where the number of pixel mutual information is decreased slightly and the number edge mutual information goes up.

Table 2 Fuse result of IR image and visual image

Evaluation	PMI	EMI
DW	0.3036	0.3931
LP	0.3105	0.4847
DWF	0.2441	0.4315
Proposed algorithm	0.3030	0.5004

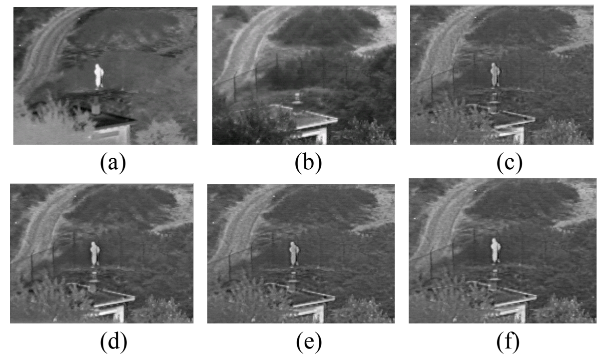

Fig.6 Fusion result of IR image and visual image. (a) Infrared image; (b) CCD image, (c) Zhang Zhong's algorithm; (d) G. Piella's algorithm; (e) Discrete wavelet frame; (f) Algorithm proposed in this paper

Fig.7 is CT and MRI image fusion image fusion. The region feature is defined by the CT image. Their fusion result can be found in Table 3.

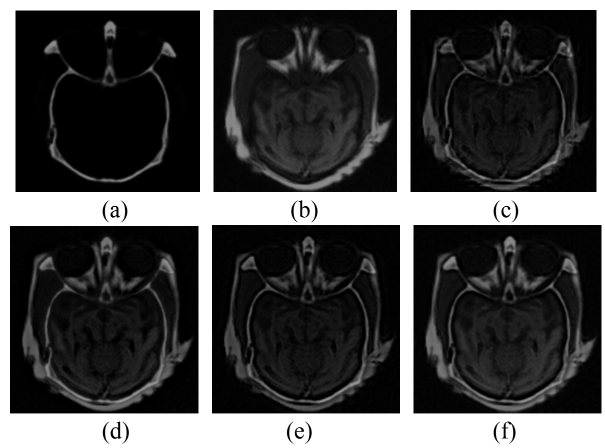

Fig.7 Fusion result of CT image and MRI image. (a) CT image; (b) MRI image; (c) Zhang Zhong's algorithm; (d) G. Piella's algorithm; (e) Discrete wavelet frame; (f) Algorithm proposed in this paper

Table 3 Fuse result of CT image and MRI image

Evaluation	PMI	EMI
DW	0.3322	0.5458
LP	0.1474	0.7468
DWF	0.4214	0.6794
Proposed algorithm	0.2830	0.7690

In Tables 1~3, DW, LP and DWF denote Zhang Zhong's wavelet based algorithm, G. Piella's Laplacian based algorithm and discrete wavelet frame based algorithm (Unser, 1995) respectively.

CONCLUSION

In this paper, an image fusion method is proposed based on fuzzification of region feature and discrete wavelet frame for merging multiple sensor images. The algorithm preserves the image contrast and obtains better region similarity than existing algorithm. Experimental results indicate that the proposed algorithm outperforms the discrete wavelet frame transform based algorithm and existing region based fusion algorithm. The feature of region and function of membership are important for improving the fusion result, so they should be chosen by considering application and the characteristics of the imaging sensor.

References

Hall, D.L., Llinas, J., 1997. An introduction to multisensor data fusion. *Proceedings of the IEEE*, **85**(1):6-23. [doi:10.1109/5.554205]

Piella, G., 2002. A General Framework for Multiresolution Image Fusion: From Pixels to Regions. Technical Report PNA-R0211, ISSN 1386-3711, CWI, Amsterdam, the Netherlands.

Qu, G., Zhang, D., Yan, P., 2002. Information measure for performance of image fusion. *Electronics Letters*, **38**(7):313-315. [doi:10.1049/el:20020212]

Rockinger, O., Fechner, T., 1998. Pixel-level image fusion: The case of image sequences. *Proceedings of SPIE—The International Society for Optical Engineering*, **3374**:378-388.

Unser, M., 1995. Texture classification and segmentation using wavelet frames. *IEEE Transactions on Image Processing*, **4**(11):1549-1560. [doi:10.1109/83.469936]

Xydeas, C.S., Petrovic, V., 2000. Objective image fusion performance measure. *Electronics Letters*, **36**(4):308-309. [doi:10.1049/el:20000267]

Zhang, Z., Blum, R.S., 1997. A Region-based Image Fusion Scheme for Concealed Weapon Detection. Proc. 31st Annu. Conf. Information Sciences and Systems, Baltimore, MD, p.168-173.

Zhou, J., Civco, D.L., Silander, J.A., 1998. A wavelet transform method to merge Landsat™ and SPOT panchromatic data. *International Journal of Remote Sensing*, **19**(4):743-757. [doi:10.1080/014311698215973]