Fusion of urban remote image based on multi-characteristics

Xuhong Yang (杨旭红)[1], Zhongliang Jing (敬忠良)[2], and Gang Liu (刘 刚)[2]

[1]*Institute of Aerospace Information and Control, School of Electronic, Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200030*

[2]*Institute of Aerospace Science & Technology, Shanghai Jiao Tong University, Shanghai 200030*

Received September 12, 2005

A fusion approach is proposed to refine the resolution of urban multi-spectral images using the corresponding high-resolution panchromatic (PAN) images. Firstly, the two images are decomposed by wavelet transformation, and five texture features are extracted from high-frequency detailed sub-images. Then a multi-characteristics fusion rule is used to merge wavelet coefficients from the two images according to the extracted features. Experimental results indicate that, comparing with the non-characteristic methods, the proposed method can efficiently preserve the spectral information while improving the spatial resolution of the urban remote sensing images.

OCIS codes: 280.0280, 330.6100, 330.6180, 350.2660.

Urban districts always have many streets, blocks, and different man-made objects, and their imageries are differentiated from imagery of non-urban regions by texture characteristic. Data fusion technique is often used to take advantages of multi-spectral data and panchromatic (PAN) data for producing spatially enhanced multi-spectral observations[1]. Along with the evolution of fusion schemes, the fusion rules have been advanced from the early simple pixel maximum or minimum selection into area-based fusion rules[2].

In this paper, five textural features are used for the representation of the characteristics of remote sensing image in urban districts.

Fractal Dimension: Fractals are very useful in modeling roughness and self-similarity properties of surfaces in image processing. The fractal dimensions of image $f(x)$ can be calculated as[2]

$$\ln E\left|f(x+\Delta x)-f(x)\right|^2 = 2H\ln\|\Delta x\| + \ln C, \qquad (1)$$

where C is a constant. The least squares solution of Eq. (1) can be got as a linearity, and its slope will be H. The fractal dimension FD can be denoted as

$$\mathrm{FD} = 3 - H. \qquad (2)$$

In order to analysis the image data in urban regions meticulously, the concept of sub-band fractal dimensions is proposed: including vertical fractal dimension FD_1, horizontal fractal dimension FD_2, and diagonal fractal dimension FD_3.

In the vertical detailed sub-image $D_j^1 f$ of size $M_1 \times N_1$, according to vertical direction, set Δx to be $k \in Z^+$ in Eq. (1). If only vertical direction is considered, the fractal Brownian function F_1 of the vertical detailed sub-image can be calculated as

$$F_1(k) = \frac{\sum\limits_{x=0}^{M_1-k-1}\sum\limits_{y=0}^{N_1-1}\left|D_j^1 f(x,y) - D_j^1 f(x+k,y)\right|^2}{N_1(M_1-k)}. \qquad (3)$$

In the horizontal detailed son-image $D_j^2 f$ of size $M_1 \times N_1$, according to horizontal direction, set Δx to be $k \in Z^+$ in Eq. (1). If only vertical direction is considered, the term of $E\left|f(x+\Delta x)-f(x)\right|^2$ in Eq. (1) should be rewritten as

$$F_2(k) = \frac{\sum\limits_{x=0}^{M_1-1}\sum\limits_{y=0}^{N_1-k-1}\left|D_j^2 f(x,y) - D_j^2 f(x,y+k)\right|^2}{M_1(N_1-k)}. \qquad (4)$$

In the diagonal detailed son-image $D_j^3 f$ of size $M_1 \times N_1$, according to diagonal direction, set Δx to be $k \in Z^+$ in Eq. (1). So the term $E\left|f(x+\Delta x)-f(x)\right|^2$ in Eq. (1) should be rewritten as

$$F_3(k) = \frac{\sum\limits_{x=0}^{M_1-k-1}\sum\limits_{y=0}^{N_1-k-1}\left|D_j^3 f(x,y) - D_j^3 f(x+k,y+k)\right|^2}{2(M_1-k)(N_1-k)} + \frac{\sum\limits_{x=0}^{M_1-k-1}\sum\limits_{y=k}^{N_1-1}\left|D_j^3 f(x,y) - D_j^3 f(x+k,y-k)\right|^2}{2(M_1-k)(N_1-k)}. \qquad (5)$$

For a non-characteristic fusion rule of fractal dimension, the minimal fractal dimension (MFD) fusion rule can be adopted[2], the data pixels whose fractal dimensions would be minimal are stored for the use of fused image.

Directive Contrast: Local contrast is used to denote the differences of objects in an image. Directive contrast includes the high frequency information of an image and the relative intensity of high frequency to the background. The directive contrast of image $f(x,y)$ at resolution 2^j can be defined as

$$\left.\begin{array}{ll} \text{Vertical Contrast} & C_j^1 = D_j^1 f / A_j f \\ \text{Horizontal Contrast} & C_j^2 = D_j^2 f / A_j f \\ \text{Diagonal Contrast} & C_j^3 = D_j^3 f / A_j f \end{array}\right\}. \qquad (6)$$

For a non-characteristic fusion rule of the directive contrast, the maximum directive contrast fusion rule (C-Max rule) can be adopted: the pixels of the source image whose directive gradient is maximum are saved as the pixels of the fused image.

Energy[3]: Energy is a parameter to measure the textural uniformity of image information. The local directive energy of image $f(x,y)$ at resolution 2^j can be defined as (for the mask of $m \times n$)

$$\left.\begin{array}{ll} \text{Vertical Energy} & E_j^1 = \sum\limits_{x=1}^{m}\sum\limits_{y=1}^{n}\left(D_j^1 f(x,y)\right)^2 \\ \text{Horizontal Energy} & E_j^2 = \sum\limits_{x=1}^{m}\sum\limits_{y=1}^{n}\left(D_j^2 f(x,y)\right)^2 \\ \text{Diagonal Energy} & E_j^3 = \sum\limits_{x=1}^{m}\sum\limits_{y=1}^{n}\left(D_j^3 f(x,y)\right)^2 \end{array}\right\}. \quad (7)$$

For a non-characteristic fusion rule of the directive energy, the maximum directive energy fusion rule (E-Max rule) can be adopted: the pixels of the source image whose directive energy is maximum will be saved for the use of fused image.

Standard Deviation: The standard deviation denotes the dispersed distribution of an image[2]. It is relative to a mean of the original image[4]. The directive standard deviation of image $f(x,y)$ at resolution 2^j can be defined as (for the mask of 3×3)

$$\left.\begin{array}{l} \text{Vertical Standard Deviation} \\ \mathrm{SD}_j^1 = \frac{1}{9}\sum\limits_{m=-1}^{1}\sum\limits_{n=-1}^{1}\left(D_j^1 f(x+m,y+n) - \bar{D}_j^1 f(x,y)\right)^2 \\ \text{Horizontal Standard Deviation} \\ \mathrm{SD}_j^2 = \frac{1}{9}\sum\limits_{m=-1}^{1}\sum\limits_{n=-1}^{1}\left(D_j^2 f(x+m,y+n) - \bar{D}_j^2 f(x,y)\right)^2 \\ \text{Diagonal Standard Deviation} \\ \mathrm{SD}_j^3 = \frac{1}{9}\sum\limits_{m=-1}^{1}\sum\limits_{n=-1}^{1}\left(D_j^3 f(x+m,y+n) - \bar{D}_j^3 f(x,y)\right)^2 \end{array}\right\}, \quad (8)$$

where $\bar{D}_j^1 f$, $\bar{D}_j^2 f$, and $\bar{D}_j^3 f$ represent the vertical, horizontal, and diagonal local mean gray scales of sub-bands of image, respectively.

For a non-characteristic fusion rule of the directive standard deviation, the maximum directive standard deviation fusion rule (D-Max rule) can be adopted: the pixels of the source image whose directive standard deviation is maximum may be stored for the fusion operation.

Gradient: Gradient can be used to measure the spatial resolution of an image. For an image $f(x,y)$ of size $M \times N$ at resolution 2^j, the local gradient is defined as

$$\left.\begin{array}{l} \text{Vertical Gradient} \\ G_j^1(x,y) = \frac{f(x,y)-f(x,\Delta y+y)}{\partial y} \\ \text{Horizontal Gradient} \\ G_j^2(x,y) = \frac{f(x,y)-f(\Delta x+x,y)}{\partial x} \\ \text{Diagonal Gradient} \\ G_j^3(x,y) = \frac{f(x,y)-f(\Delta x+x,\Delta y+y)}{\sqrt{(\partial x)^2+(\partial y)^2}} \end{array}\right\}. \quad (9)$$

For a non-characteristic fusion rule of the directive gradient, the maximum directive gradient fusion rule (G-Max rule) can be adopted: the pixels of the source image whose directive gradient is maximum are saved as the fused image.

In this paper, a multi-characteristics fusion rule (MC rule) is adopted for the fusion of five extracted features. In order to take advantages of the five extracted features, a fuzzy function is applied in the process of fusion. Figure 1 shows the fuzzy function.

For the original image $A(x,y)$, $B(x,y)$ of size $M \times N$ and the fusion result $P(x,y)$, the whole rule can be illustrated by

$$P(x,y) = \sum_{k=1}^{3}\omega_k \times P_k(x,y). \quad (10)$$

Firstly, the five features are normalized. Then the fusion process can be done as

$$\text{norm} = [(\mathrm{FD}_A - \mathrm{FD}_B)^2 + (C_A - C_B)^2 + (E_A - E_B)^2 + (\mathrm{SD}_A - \mathrm{SD}_B)^2 + (G_A - G_B)^2]^{1/2}, \quad (11)$$

where $0 \leq \text{norm} \leq 5$.

For ω_1, the fusion can be calculated as

$$D_j^i f_{P_1}(x,y) = \left(D_j^i f_A(x,y) + D_j^i f_B(x,y)\right)/2. \quad (12)$$

For ω_2, five counters are used

$$\begin{cases} \text{count1}_j^i = 1, & \text{if } \mathrm{FD}_{j,A}^i(x,y) \leq \mathrm{FD}_{j,B}^i(x,y) \\ \text{count1}_j^i = 0, & \text{else} \end{cases},$$

$$\begin{cases} \text{count2}_j^i = 1, & \text{if } C_{j,A}^i(x,y) \geq C_{j,B}^i(x,y) \\ \text{count2}_j^i = 0, & \text{else} \end{cases},$$

$$\begin{cases} \text{count3}_j^i = 1, & \text{if } E_{j,A}^i(x,y) \geq E_{j,B}^i(x,y) \\ \text{count3}_j^i = 0, & \text{else} \end{cases},$$

$$\begin{cases} \text{count4}_j^i = 1, & \text{if } \mathrm{SD}_{j,A}^i(x,y) \geq \mathrm{SD}_{j,B}^i(x,y) \\ \text{count4}_j^i = 0, & \text{else} \end{cases},$$

$$\begin{cases} \text{count5}_j^i = 1, & \text{if } G_{j,A}^i(x,y) \geq G_{j,B}^i(x,y) \\ \text{count1}_j^i = 0, & \text{else} \end{cases},$$

$$i = 1,2,3 \quad \text{at resolution } 2^j, \quad (13)$$

Then the fusion result can be calculated as

$$\text{sum}_j^i = \text{count1}_j^i + \text{count2}_j^i + \text{count3}_j^i + \text{count4}_j^i + \text{count5}_j^i,$$

$$\begin{cases} D_j^i f_{P_2}(x,y) = D_j^i f_A(x,y), & \text{if } \text{sum}_j^i \geq 3 \\ D_j^i f_{P_2}(x,y) = D_j^i f_B(x,y), & \text{else} \end{cases},$$

$$\text{at resolution } 2^j, \quad (14)$$

$i = 1,2,3$ for vertical direction, horizontal direction, and diagonal direction, respectively.

It follows that the biggest average gradient may appear at the fused image by means of C-Max fusion rule. So, for ω_3, the fusion can be calculated as

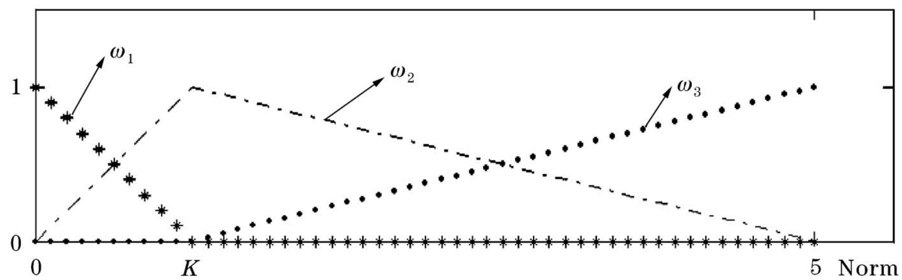

Fig. 1. Fuzzy function.

Table 1. Evaluation Results of the 1st Group Images

	CC (3 Tunnels of RGB)			DI (3 Tunnels of RGB)			AG (3 Tunnels of RGB)		
	R	G	B	R	G	B	R	G	B
MFD	0.93692	0.93259	0.93641	0.65121	0.54734	0.74654	11.577	11.617	11.533
E-Max	0.88422	0.8767	0.87927	0.72474	0.6245	0.8233	17.049	17.036	17.019
C-Max	0.88789	0.88062	0.88316	0.71269	0.6159	0.81266	17.207	17.213	17.195
D-Max	0.88422	0.8767	0.87928	0.72492	0.62475	0.82339	17.046	17.033	17.017
G-Max	0.88689	0.87953	0.88215	0.72011	0.62016	0.81808	17.064	17.071	17.053
MC	**0.89349**	**0.88574**	**0.88748**	**0.49982**	**0.46649**	**0.56274**	**17.343**	**17.312**	**17.315**

$$D_j^i f_{P_3}(x,y) = \begin{cases} D_j^i f_A(x,y), \text{ if } C_{j,A}^i(x,y) \geq C_{j,B}^i(x,y) \\ D_j^i f_B(x,y), \text{ else} \end{cases}, \quad \text{at resolution } 2^j, \tag{15}$$

$i = 1, 2, 3$ for vertical direction, horizontal direction, and diagonal direction, respectively.

The fusion rule in Ref. [5] is taken to the fusion of low frequency parts. In the process of experiment it is found that this fusion rule has a significant effect on fusion result when applying to low-frequency parts.

The proposed algorithm consists of the following steps:

1) Transform the R, G, and B bands of the multi-spectral image into intensity-hue-saturation (IHS) format, then the PAN image and the intensity component of the multi-spectral image are decomposed into a wavelet representation at the same coarser resolution.

2) Calculation five features of wavelet transform coefficients as stated before.

3) Fusion based on the MC fusion rule.

4) Inversely discrete wavelet transforms (IDWTs) and inversely IHS transform are performed to get the merged RGB image.

Three assessment criteria are used in this paper.

Correlation Coefficient (CC): The CC between the original image and the merged image can represent their correlativity, and can be calculated as in Ref. [5].

Deviation Index (DI): An image's spectrum DI may indicate the distortion of the multi-spectral image

$$\mathrm{DI} = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left| \frac{P(x,y) - A(x,y)}{A(x,y)} \right|. \tag{16}$$

Average Gradient (AG): An image's AG may express the clarity of an image. For an image $P(x,y)$ of $M \times N$, the AG is defined as

$$\mathrm{AG} = \frac{1}{(M-1)\times(N-1)} \sum_{x=1}^{(M-1)} \sum_{y=1}^{(N-1)} \sqrt{\left[\left(\frac{\partial P(x,y)}{\partial x}\right)^2 + \left(\frac{\partial P(x,y)}{\partial y}\right)^2\right] \Big/ 2}. \tag{17}$$

All these fusion methods are used to merge the Quick-Bird PAN image in 0.61-m resolution and the corresponding Quick-Bird multi-spectral image in 2.44-m resolution at the Shanghai district in 2002. The size of images is 512×512.

Table 1 shows the evaluation results of correlation coefficient (CC), deviation index (DI), and average gradients (AG). The CC of the proposed method is higher than the others and a little lower than MFD rule, but still can be accepted. The DI of the proposed method is the smallest among the fusion rules. The AG of the proposed method is much bigger than the other methods. The assessment criteria indicate that the proposed method can refine the multi-spectral images more clearly than the other methods without introducing extra spectral distortion.

Figure 2 shows fusion results by use of different fusion rules. Figures 2(a) and (b) show the original images of Quick-Bird PAN image in 0.61 m and the corresponding multi-spectral images in 2.44 m. Fusion results by use of different fusion rules are shown in Figs. 2(c)—(h) respectively. It can be seen from the fusion images that the multi-spectral images are refined and may well preserve some spectral information.

The conclusion can be stated as below: the proposed algorithm aims at sharpening urban multi-spectral information from corresponding high-resolution panchromatic

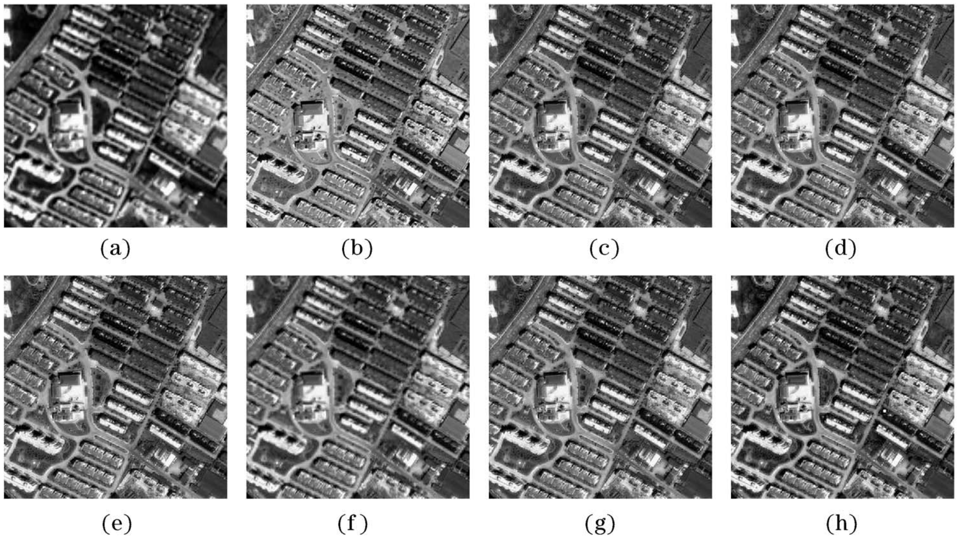

Fig. 2. Fusion results by use of different fusion rules. (a) Original multi-spectral image; (b) original panchromatic image; (c) fusion result of MFD (fractal minimum fusion rule); (d) fusion result of E-Max (energy maximum fusion rule); (e) fusion result of C-Max (contrast maximum fusion rule); (f) fusion result of D-Max (standard deviation maximum fusion rule); (g) fusion result of G-Max (gradient maximum fusion rule); (h) fusion result of MC (promoted method).

images. The multi-characteristics fusion rule is used to combine five kinds of features extracted from the original images. The proposed method can efficiently preserve the spectral information and improve the spatial resolution of remote sensing images.

Image fusion using high spatial resolution PAN data and multi-spectral data can give better result in details and useful spectral information, which in turn can discriminate small objects or land cover types. Fused images are primarily used: (a) to be presented to human observers for viewing or interpretation; (b) to be further processed by a computer using different image processing techniques.

This work was supported by the National "863" Projects (No. 2001AA135091), the Shanghai Key Project (No. 02DZ15001), and the China Aviation Science Foundation (No. 02D57003). X. Yang's e-mail address is yxh@sjtu.edu.cn or lyqq02@163.com.

References

1. L. Alparone, A. Garzelli, F. Nencini, B. Aiazzi, and S. Baronti, Proc. SPIE **5238,** 44 (2004).
2. J. Tian, J. Chen, and C. Zhang, Proc. SPIE **5308,** 824 (2004).
3. D. I. Morales, M. Moctezuma, and F. Parmiggiani, in *Proceeding of IGARSS '03* **6,** 3504 (2003).
4. M. González-Audícana, J. Saleta, R. García Catalán, and R. García, IEEE Transactions on Geoscience and Remote Sensing **42,** 1291 (2004).
5. F. Berizzi, M. Martorella, G. Bertini, A. Garzelli, F. Nencini, F. Dell'Acqua, and P. Gamba, in *Proceedings of IGARSS'04* **1,** 93 (2004).