A Variational Gradient-based Fusion Method for Visible and SWIR Imagery

Huifang Li, Liangpei Zhang, Huanfeng Shen, and Pingxiang Li

Abstract

This paper presents a new variational gradient-based fusion method for visible and short-wave infrared (SWIR) imagery. The proposed method enables spatial enhancement and dehazing of visible imagery. Integrating gradients from SWIR imagery into visible imagery produces a single image with true color and sharp gradients. A constraint based on band correlation is included to improve the enhancement and implement dehazing. The band correlation is according to the quantitative relationship between the wavelength and the atmospheric effect caused by Rayleigh scattering. In this study, both clear and hazy Landsat ETM+ images are used in the experiments. By visual assessment, the gradient of the fused image is more salient than that of the original image, and the true color is well preserved. With the inclusion of the band correlation constraint, the proposed fusion method yields almost haze-free results. Quantitatively, the Metric Q of the fused images is significantly higher than that of the original images; the largest increase of the Metric Q in the experimental results is from 0.0114 to 0.0611. Moreover, for the results of the proposed method, the Metric Q increase in the visible bands declines from blue band to red band.

Introduction

Atmospheric scattering changes the spatial distribution of energy and causes degradation problems with images captured by sensors. This scattering depends on the wavelength of the light and the particle size in the atmosphere (Nayar *et al.*, 1999). When the particle size is much smaller than the wavelength, Rayleigh scattering is dominant (Jacobsen *et al.*, 2000; Su *et al.*, 2006). Clear air is composed of molecules, whereas haze is comprised of an aerosol, which is a dispersed system of small particles suspended in gas (Nayar *et al.*, 1999). Therefore, on a clear or hazy day, contaminations in visible imagery, such as blurring and hazing, can be mainly attributed to the Rayleigh scattering. Due to its long wavelength, infrared imagery is free from haze degradation and endowed with sharp gradients. Hence, the proposed method corrects for Rayleigh scattering in visible imagery by fusing spatial geometry information from the infrared imagery.

The proposed correction method is a kind of image-based correction versus model-based correction. Image-based correction is independent of *in situ* measurements of the atmosphere, the sun and the sensor; therefore, its applicability is broad (Chavez, 1996). A number of image-based correction methods have been previously used, and we review them here briefly. Dark object subtraction (DOS) assumes that dark objects have zero radiance (Norjamaki *et al.*, 2007). Consequently, the radiance of dark objects responds to the atmospheric path radiance, and it is then subtracted from the original pixel values. When applied in dehazing, DOS is based on haze detection, which distinguishes hazy and non-hazy regions. Tasseled cap (TC) (Crist *et al.*, 1984), threshold control (Richter, 1996a; Richter, 1996b) and manual drawing (Liang, 2001) are common methods of haze detection. Zhang developed haze optimized transformation (HOT) to detect hazy regions in Landsat images, and then DOS based on HOT was employed for haze removal (Zhang *et al.*, 2002; Zhang *et al.*, 2003). Considering multiplicative effects, He *et al.* (2010) proposed a virtual cloud point (VCP) method based on advanced HOT (AHOT) to remove haze. A classical image processing method, histogram matching (HM), has also been used for dehazing. As well as DOS and VCP, HM needs haze-free regions as references before correcting a single-band image (Stark, 2000; Liang, 2001). The histogram of hazy regions is then matched to that of the clear regions. Hence, precise haze detection is essential before haze removal in DOS, VHP, and HM; if clear regions are absent, these methods fail. Besides, the above methods are designed for single-band imagery. For multispectral imagery they must be operated band-by-band. Multispectral imagery consisting of several bands has abundant spectral information, and different bands have different degrees of atmospheric contaminates. Ignorance of the band correlation in multispectral imagery usually causes spectral distortion. Therefore, an automatic correction method for multispectral imagery must involve band correlation. Once the hazy channels are dehazed, the spectrum of each pixel in multispectral imagery is restored. It may enable an increase in the accuracy of applications such as classification, target detection, change detection, etc., which usually involve multi-spectral calculation (Huang *et al.*, 2007; Huang *et al.*, 2008; Zhang *et al.*, 2010).

Gradient-based image fusion is composed of three steps: first, gradients are extracted from the input images; then, the target gradient is defined by fusion rules, which are manipulated at the feature level; lastly, the final output is

Huifang Li, Liangpei Zhang, and Pingxiang Li are with The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129, Luoyu Rd., Hong Shan, Wuhan, China (zlp62@public.wh.hb.cn).

Huanfeng Shen is with the School of Resource and Environmental Science, Wuhan University, 129, Luoyu Rd., Hong Shan, Wuhan, China.

Photogrammetric Engineering & Remote Sensing
Vol. 78, No. 9, September 2012, pp. 000–000.

0099-1112/12/7809–0000/$3.00/0

reconstructed from the given target gradient at the signal level. In general, gradient-based fusion is a kind of signal-level fusion; it assumes that the most important input information is faithfully represented in the fused image at the signal level (Petrovic *et al.*, 2004). Gradient-based fusion has been proven to be effective in reducing reconstruction error artifacts and preserving contrast information. Different fusion rules at the feature level in the gradient domain have been employed by researchers to define the target gradient. The most common rules include combining gradients by a direct union, decision tree learning, and weighted averaging (Davis *et al.*, 2007; Aksoy *et al.*, 2009). The weights in weighted averaging can be determined according to the specific application (Raskar *et al.*, 2005). A simple choice, used in the pan-sharpening process, is to weight panchromatic and multispectral images as 1 and 0 (Ballester *et al.*, 2006; Zhou *et al.*, 2010). Socolinsky extended the definition of gradient for single-band images to multi-band images through first-order fusion (Socolinsky *et al.*, 1999; Socolinsky *et al.*, 2002). Piella did the related work using a structure tensor to obtain the fused gradient (Piella, 2009). The above-mentioned techniques have been used to improve image geometry and visual quality. The variational framework has been frequently adopted in these studies due to its flexibility and extensibility. It enables terms corresponding to natural physical conditions (such as the reflectance is nonnegative, the lightness is spatially smooth, etc.) to be integrated into the framework. Three terms are contained in P+XS image fusion (Ballester *et al.*, 2006), four in perceptually enhanced fusion (Piella, 2009), and five in fast variational fusion for pan-sharpening (Zhou *et al.*, 2010). However, with the increasing number of terms, how to automatically determine the influence weight for each term is a problem. Therefore, realizing correction by using fewer limiting terms will improve the automation and minimize the need for human intervention. The proposed variational fusion method containing two terms is presented in the following sections.

Gradient-based Fusion (GF)

Definition

Gradient-based fusion (GF) integrates information from two or more images into a single image which is more informative than any of the input images (Burt *et al.*, 1993; Pohl *et al.*, 1998). Usually, we undertake image fusion in remote sensing applications to improve the spatial resolution, which inevitably increases the amount of data. However, gradient-based fusion aims to enrich images with large amounts of information by adjusting image contrast rather than increasing the amount of data. Therefore, results from gradient fusion do not actually have higher spatial resolution than the input images, but are more visually pleasing.

Specifically, supposing two images are taken as the input data, one has abundant color or spectral information but with poor spatial geometrical information. The other one is just a single-band image but has sharp and clear edges and textures. In order to obtain the final colorful, clear, and sharp image, the second image is considered as the reference data which provides the ideal gradients. The gradient-based fusion process is then realized by using the variational framework expressed as Equation 1, including two terms: the first term ensures that the final gradient approaches the reference gradient; the second term ensures the consistency of the final and original data. Note that this model is based on the assumption that the result gradient is linearly related to the reference gradient.

$$\min F(u) = \frac{1}{2}\iint_{\Omega} \|\nabla u - \nabla f\|^2 + \alpha \|u - u_0\|^2 dxdy \tag{1}$$

where ∇u and ∇f are gradient vectors of the final result and the reference data, respectively, which are defined as $\nabla u = [u_x \quad u_y]^T$ and $\nabla f = [f_x \quad f_y]^T$, where $u_x = u_{i,\,j+1} - u_{i,\,j}$ and $u_y = u_{i+1,\,j} - u_{i,\,j}$, u_0 is the original data, Ω is the region of the imagery, and α is a nonnegative parameter to balance the weight of the first and second terms.

Moment Matched Gradient-based Fusion (MGF)

Data from different sources are endowed with distinct radiometric resolution; 8-bit data has a radiometric resolution of 0 to 255, for example. However, not all data will take up the entire range. This is then called the dynamic range of the data. For example, a certain image in band 7 might have a dynamic range of 0 to 30. Thus, direct approximation of ∇u to ∇f is inappropriate, especially when the dynamic ranges are notably different. Making the dynamic ranges of the unknown and referenced variable equal is the first step before fusion. Since the resulting gradient ∇u has linear correlation with the reference gradient ∇f, adjusting the mean and deviation of ∇u is sufficient to achieve this goal.

Moment matching has been successfully used by researchers to remove stripes in scenes by adjusting DNs between sub-scenes (Gadallah *et al.*, 2000); it is an efficient way to regularize one image to another. Hence, moment matching of ∇u and ∇f is taken as the pretreatment to unify the dynamic range before fusion. The adjusted ∇u is expressed as $\nabla \hat{u}$, which will replace ∇u in Equation 1:

$$\nabla \hat{u} = \frac{(\nabla f)_{std}}{(\nabla u)_{std}} (\nabla u - (\nabla u)_{mean}) + (\nabla f)_{mean} \tag{2}$$

where $(\nabla u)_{mean}$ and $(\nabla f)_{mean}$ are the means of ∇u and ∇f, and $(\nabla u)_{std}$ and $(\nabla f)_{std}$ are the standard deviations of ∇u and ∇f. Therefore, the gradient-based fusion model can be updated to the moment matched gradient-based fusion (MGF), shown as follows:

$$\min F(u) = \frac{1}{2}\iint_{\Omega} \|\nabla \hat{u} - \nabla f\|^2 + \alpha \|u - u_0\|^2 dxdy \tag{3}$$

Numerical Solution

The GF model is linear because only $L2$-norm constraints, represented by the square of the Euclidean norm, are included, so that its convergence is obvious. The simplest method, steepest descent, has been employed to solve this optimal model. The solution is obtained by the following iteration:

$$u_{m+1} = u_m - tG, \quad m \geq 0 \tag{4}$$

where t is the step, and G is the gradient of F at u. The Euler-Lagrange Equation 3 can be expressed as:

$$0 = -\frac{(\nabla f)_{std}^2}{(\nabla u)_{std}^2} \cdot \Delta u + \frac{(\nabla f)_{std}}{(\nabla u)_{std}} \cdot \Delta f + \alpha(u - u_0). \tag{5}$$

Therefore:

$$G = \frac{\partial F}{\partial u} = -\frac{(\nabla f)_{std}^2}{(\nabla u)_{std}^2} \cdot \Delta u + \frac{(\nabla f)_{std}}{(\nabla u)_{std}} \cdot \Delta f + \alpha(u - u_0), \tag{6}$$

$$t = \frac{G^2}{-G\dfrac{(\nabla f)_{std}^2}{(\nabla u)_{std}^2} \cdot \Delta G + \alpha G^2} \tag{7}$$

The step and gradient are substituted into Equation 4, and the result can be obtained after numbers of iterations.

GF Model for Dehazing

It is well known that the atmospheric effect varies in channels with different wavelengths. Though the wavelengths of visible channels are close to each other, differences and correlations of atmospheric effects exist among them on both clear and hazy days. However, these differences and correlations are not considered in the MGF model.

In the optical remote sensing field, atmospheric effects on the observed intensity of the land surface are closely associated with the atmospheric optical properties. However, not all the *in situ* atmospheric parameters are available as some archived data lose their attribute data. This is the main obstacle when undertaking absolute atmospheric correction (Mahiny *et al.*, 2007; Davranche *et al.*, 2009). Image-based correction aims to correct the contaminated images to clear and informative images, without *in situ* parameters, attenuating or even eliminating the effects of the uneven atmosphere.

In terms of the radiative transfer equation, the variation of the intensity in a certain direction is twofold. On one hand, the absorbing and scattering in this direction lead to intensity reduction; on the other hand, scattering from all the other directions increases the intensity in this direction and reduces the image contrast. The sum of the intensity of multiple scattered light accounts for the final intensity in this certain direction. In a cloudless atmosphere, the contribution of multiple scattered light decreases rapidly, and thus the total intensity can be expressed as a finite sum. In cloudy conditions, the intensity is a sum of all the scattering orders, which explains why the intensity in hazy or cloudy areas is usually higher than that in clear regions. Therefore, reducing intensity and enhancing the contrast in the hazy and cloudy regions is an effective way of dehazing.

$$I(\mu,\varphi) = I_0(\mu,\varphi) - I_{a,s}(\mu,\varphi) + \sum_i I_s(\mu_i,\varphi_i) \tag{8}$$

where μ and φ are the zenith and azimuth angles of the direction of propagation of the light, (μ, φ) indicates the light propagating direction, $I(\mu, \varphi)$ is the observed intensity in the certain direction, $I_0(\mu, \varphi)$ is the supposed intensity without atmospheric influences, $I_{a,s}(\mu, \varphi)$ represents the absorbed and scattered intensity, and $\sum_i I_s(\mu_i,\varphi_i)$ records the scattered intensity from all the other directions.

The second term in the above equation causes the intensity reduction of a pixel, and the third term is the actual reason for the intensity increase. It can be seen that the intensity of a pixel is higher when it is hazy than when it is clear. Thus, intensity reduction is offset by the intensity increase caused by scattering from all the other directions. Scattering is the dominant factor for the haze rather than the absorption. Therefore, roughly, we can assume $I_{a,s}(\mu, \varphi) = 0$, so $I_{haze}(\mu,\varphi) = \sum_i I_s(\mu_i,\varphi_i)$. Hence, Equation 9 can be simplified to:

$$I(\mu,\varphi) = I_0(\mu,\varphi) + I_{haze}(\mu,\varphi). \tag{9}$$

Therefore, in order to take the band correlation into consideration, the next task is to understand the relationship between the wavelength and the intensity change caused by haze.

Relationship between the Wavelength and the Atmospheric Effect

As mentioned before, we deem that Rayleigh scattering dominates the atmospheric scattering in visible channels on both clear and hazy days. Therefore, for a particle with a certain size, suspended in the atmosphere, the intensity of the scattered light is inversely proportional to the biquadratic wavelength. Briefly, the relationship between the scattered intensity and the wavelength can be expressed as (McCartney, 1976; Bucholtz, 1995):

$$I_s(\mu_i,\varphi_i) = C_s(\mu_i,\varphi_i)\lambda^{-4} \tag{10}$$

where $C_s(\mu_i, \varphi_i)$ is a parameter associated with the particle size and the scattering angle. Figure 1 shows the curve of the scattered intensity versus the wavelength. The scattered intensity decreases with the increase in light wavelength. For the 2.2 μm channel, Rayleigh scattering is so faint that it can be ignored.

Therefore, as a sum of the intensity of the scattered light from all the other directions, the haze intensity can be expressed as:

$$I_{haze} = C\lambda^{-4} \tag{11}$$

where $C = \sum_i C_s(\mu_i, \varphi_i)$. C can be considered as a constant for a certain location corresponding to a pixel in the imagery, but it differs between locations. We cannot estimate the C for each pixel and calculate the intensity of the scattered light accurately. Therefore, adding a constraint, derived from the relationship between the haze intensity and the wavelength, into a variational function, is a practical way to solve this dehazing problem. It can be seen that, for a specific pixel, the ratio of the haze intensity in different visible bands only relates to the wavelength, i.e.,

$$I_{haze,\lambda_1} : I_{haze,\lambda_2} : \ldots : I_{haze,\lambda_K} = \lambda_1^{-4} : \lambda_2^{-4} : \ldots : \lambda_K^{-4} \tag{12}$$

where K is the total number of bands. This relative relation among the haze intensity provides us with a useful constraint. If the intensity change in each visible band agrees with this proportion, the result should avoid the haze contamination, along with the physical effects of the atmosphere. Thus, the next key step is to project the intensity change calculated from the variational function to the ratio relation as Equation 12. Taking the K bands as K samples of one pixel, we intend to find a statistic relating to the sample entirety, but independent of each single sample. Here, the mean is taken as this representation, i.e.,

$$\bar{I}_{haze} = \frac{C}{K}\sum_{i=1}^{K}\lambda_i^{-4}. \tag{13}$$

The ratio of I_{haze,λ_i} to $\bar{I}_{haze}$ is:

$$R_i = \frac{I_{haze,\lambda_i}}{\bar{I}_{haze}} = \frac{K}{\sum_{j=1}^{K}\left(\frac{\lambda_j}{\lambda_i}\right)^{-4}}. \tag{14}$$

This ratio is a constant for a given channel, which ensures the proportional relationship among bands. We take advantage of this ratio value to project the intensity change in the variational function to the proportional relationship as Equation 12. The projection map is illustrated by Figure 1. The right vertex of the average haze intensity line is the projection center. Assuming that the change intensity mean of the variational function is $\bar{I}_{change}$, as labeled in Figure 1, the vertical line perpendicular to the blue line is the projection plane. Crossing the projection center, the intensity of

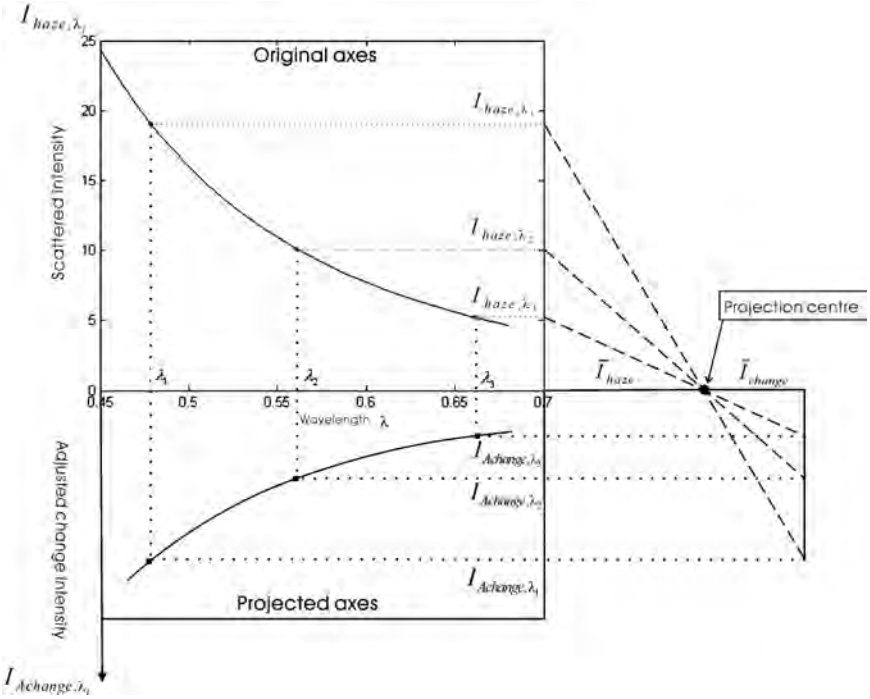

Figure 1. Mean haze projection (MHP); it adjusts the relationship of the change intensity and the wavelength to agree with the Rayleigh scattering rule.

haze is mapped to the projection plane. We then get the adjusted change intensity in the projected axes:

$$I_{Achange,\lambda_i} = R_i \bar{I}_{change}. \tag{15}$$

It is easy to prove that the adjusted change intensities satisfy the exponential relationship and also the proportional relation as Equation 12 is fulfilled, i.e.,

$$I_{Achange,\lambda_1} : I_{Achange,\lambda_2} : \ldots : I_{Achange,\lambda_k} = \lambda_1^{-4} : \lambda_2^{-4} : \ldots : \lambda_K^{-4}. \tag{16}$$

Therefore, the exponential curve can fit the three points in the new projected axes well, as shown in Figure 1. We define this central projection as the mean haze projection (MHP). Moreover, since the imaging wavelengths are arranged from short to long sequentially, thus $R_1 > R_2 > \ldots R_K$ and $R_{\max} = R_1 > 1$, $R_{\min} = R_K < 1$, $R_i \in [R_1, R_K]$. Therefore, the adjusted change intensity increases monotonically with the wavelength, i.e., $I_{Achange,\lambda_1} > I_{Achange,\lambda_2} > \ldots > I_{Achange,\lambda_k}$, which is consistent with light scattering theory.

Constrained MGF (CMGF)

Based on the above definition and analysis of the mean haze projection (MHP), we propose the constrained MGF (CMGF) model, taking the MHP as a constraint in the variational MGF model. The MHP is based on band correlation and is used to control the iterative process of the variational MGF model. The CMGF aims to suppress the atmospheric scattering effects and remove the slight haze while avoiding spectral distortion. As mentioned before, the numerical solution of the variational MGF is obtained by the iteration Equation 4, so that tG corresponds to the change intensity in each iteration I_{change,λ_i}, i.e., $I_{change,\lambda_i} = tG$. If the MHP constraint is implemented in each iteration, the final adjusted change intensity will fulfill the negative exponent relationship. Consequently, the CMGF can be expressed as:

$$\min F(u) = \frac{1}{2} \iint_{\Omega} \| \nabla \hat{u} - \nabla f \|^2 + \alpha \| u - u_0 \|^2 dxdy$$

$$\text{with the MHP: } u_{change} \rightarrow u_{Achange} \tag{17}$$

where $u_{change} = tG$, and $u_{Achange}$ represents the adjusted u_{change} and $u_{Achange,\lambda_i} = R_i \bar{u}_{change}$. The iteration equation is:

$$u_{m+1} = u_m - u_{Achange}. \tag{18}$$

Taking the MHP in the iteration has the effect of imposing a forcible constraint in the variational model. It is noted that all the bands of the multispectral image u are solved simultaneously in the CMGF method, rather than calculating band-by-band as in the MGF. Also, the correlation among different bands is considered in the solving process. Therefore, spectral distortion can be minimized or even avoided in the results. The convergence of the CMGF will be discussed in the experimental section.

GF for Landsat ETM + images

Landsat ETM+ images consist of eight spectral bands with a spatial resolution of 30 m for bands 1 to 7, and 15 m for band 8 (panchromatic). The spectral window of the sensor ranges from 0.45 μm to 12.50 μm and includes visible, near-infrared, SWIR, and long-wave infrared bands. Contamination caused by the atmosphere in the scenes varies in the different channels because atmospheric scattering is highly wavelength dependent (Wu *et al.*, 2005); this causes the land surface features to be presented in different ways. The combination of visible

bands can present a scene with true color, which satisfies the human visual perception; however, this scene is usually hazy or cloudy because visible channels are sensitive to atmospheric conditions. In contrast, the NIR and SWIR channels with long wavelengths are insensitive to the semi-transparent atmosphere, so that scenes in these bands are usually free of haze and clouds.

Visible bands of ETM+ data, including bands 1 to 3 with low contrast, are the initial inputs in the variational gradient fusion model. The next problem is how to select a reference gradient data from three of the infrared channels, including band 4, 5, and 7. Given the assumption of gradient linear correlation, the task is to find one channel whose image gradient has the highest linear correlation or similarity to that of the visible channels. Since gradient calculation is a linear operator, the problem comes down to searching for the channel most highly related to the visible channels. Images taken in the near-infrared band 4 (0.76 to 0.90 μm) usually show opposite contrast or gradient to that of the visible channels, owing to the high reflection of vegetation and high absorption of water. In his study of the aerosol-free vegetation index (AFRI), Karnieli *et al.* (2006) demonstrated that $AFRI_{2.1}$ performs better than $AFRI_{1.6}$, where $AFRI_{2.1} = (\rho_{NIR} - 0.5\ \rho_{2.1})/(\rho_{NIR} + 0.5\ \rho_{2.1})$ and $AFRI_{1.6} = (\rho_{NIR} - 0.66\ \rho_{1.6})/(\rho_{NIR} + 0.66\ \rho_{1.6})$. Taking ETM+ data, the grey level correlation coefficients are calculated and shown in Figure 2. It can be seen that high self-correlation exists in the three visible bands, and band 7 has the highest correlation with visible bands among all the infrared bands.

The SWIR (2.1 to 2.2 μm) channel in remote sensing has been implemented in vegetation detection (Kaufman, 1994; Karnieli *et al.*, 2001) and aerosol depth estimation (Fallah-Adl *et al.*, 1996; Fallah-Adl *et al.*, 1997) due to its two special characteristics:

1. The wavelength is long enough to penetrate the atmosphere, including that containing amounts of aerosols such as smoke. Thus, scenes imaged in the SWIR channel show clear and sharp edge and texture information, avoiding the atmospheric effects.
2. Even though reflection varies in different spectral bands, linear correlation exists between visible and SWIR bands. Especially for vegetation, reflection reduction caused by chlorophyll absorption in the red band is consistent with the same phenomenon caused by liquid water absorption in plants in the SWIR band.

The first characteristic ensures that scenes in SWIR are haze/cloud free and reflects their potential assistant effects for haze/cloud removal of visible scenes. The second characteristic was studied and demonstrated by Kaufman, and the linear correlations were expressed as $\rho_{0.49} = 0.25\ \rho_{2.1}$ and $\rho_{0.66} = 0.5\ \rho_{2.1}$ (Kaufman *et al.*, 1997). Various land surface types were employed to verify these linear relations, of which the highest fitting error (around 0.03) appeared in urban and soil areas. Karnieli supplemented the relationship of 0.55 μm and 2.1 μm channels, written as $\rho_{0.55} = 0.33\ \rho_{2.1}$, and developed a new haze/cloud independent vegetation index, AFRI, using the 2.1 μm channel (Karnieli *et al.*, 2001). He assumed that the linear correlation was global and could be used for different land surface types. Vegetation and bare soil were involved in this study, the reflectance of which ranged from 0 to 0.8.

Considering the previous achievements and the analysis, we chose to take band 7 of the ETM+ data as the ancillary data supplying source gradients in the gradient-based fusion method.

	B1	B2	B3	B4	B5	B7
B1	1	0.67	0.79	-0.19	0.22	0.46
B2	0.67	1	0.67	0.43	0.47	0.49
B3	0.79	0.67	1	0.11	0.49	0.65
B4	-0.19	0.43	0.11	1	0.64	0.32
B5	0.22	0.47	0.49	0.64	1	0.91
B7	0.46	0.49	0.65	0.32	0.91	1

Figure 2. Grey level correlation coefficient matrix of ETM + data.

Results and Discussion

The gradient-based fusion improves image quality by enriching the information of edges and textures. To test its validation, five Landsat ETM+ data were selected, encompassing forest, water, urban, farmland, and soil land surface types. The data were captured on two different dates: 19 March 2002, with a clear sky, and 13 October 2002, with the sky partially covered by haze.

Clear Data

On a day with clear sky, the atmosphere is almost transparent and the main scattering is Rayleigh scattering, which reduces the image contrast evenly in the spatial domain. Results of MGF and CMGF are compared and shown in Figure 3 and Figure 4.

Three types of land surface: forest, water, and urban areas, are contained in Figure 3. In the original image, the detailed textures are blurred and vague, whereas they are enhanced and clear in the results of MGF and CMGF, as shown in Figure 3b and 3c, and the detailed regions cropped from 3b and 3c. This can be attributed to the integration of the spatial information extracted from band 7. The divisions in water regions caused by different suspension and sediment distribution are weakened in the result of MGF because water with different qualities is treated equally in band 7. However, these divisions are retained well in the results of CMGF, thanks to the MHP constraint. In the urban areas, because of the existence of abundant edges and textures, over-enhancement appears in the result of MGF, as shown in Figure 3h. All the edges, such as roads and the transitions between two regions, are highlighted in the fused results, whereas some of them are neglected in the original scenes. The intensity of urban areas is usually higher than the other regions; consequently, the amount of high values appearing in urban areas leads to over-enhanced results. However, CMGF controls the over-enhancement efficiently by adding the MHP constraint based on the band correlation. This constraint obeys the physical mechanism of the atmospheric scattering so that it ensures the high similarity of the resulting spectra and the true spectra. The contrast between pixels is retained while the edge information is enhanced, as shown in Figure 3i.

To evaluate the GF results, the basic statistics, including the means and variances of the different channels in an image, are shown in Table 1. The means of the two fused images are approximately equivalent to the original mean, which ensures the basic color consistency. The variance of the result of MGF declines when compared with the original variance, which is due to the existence of large areas of water surface and the deflation of the result data range.

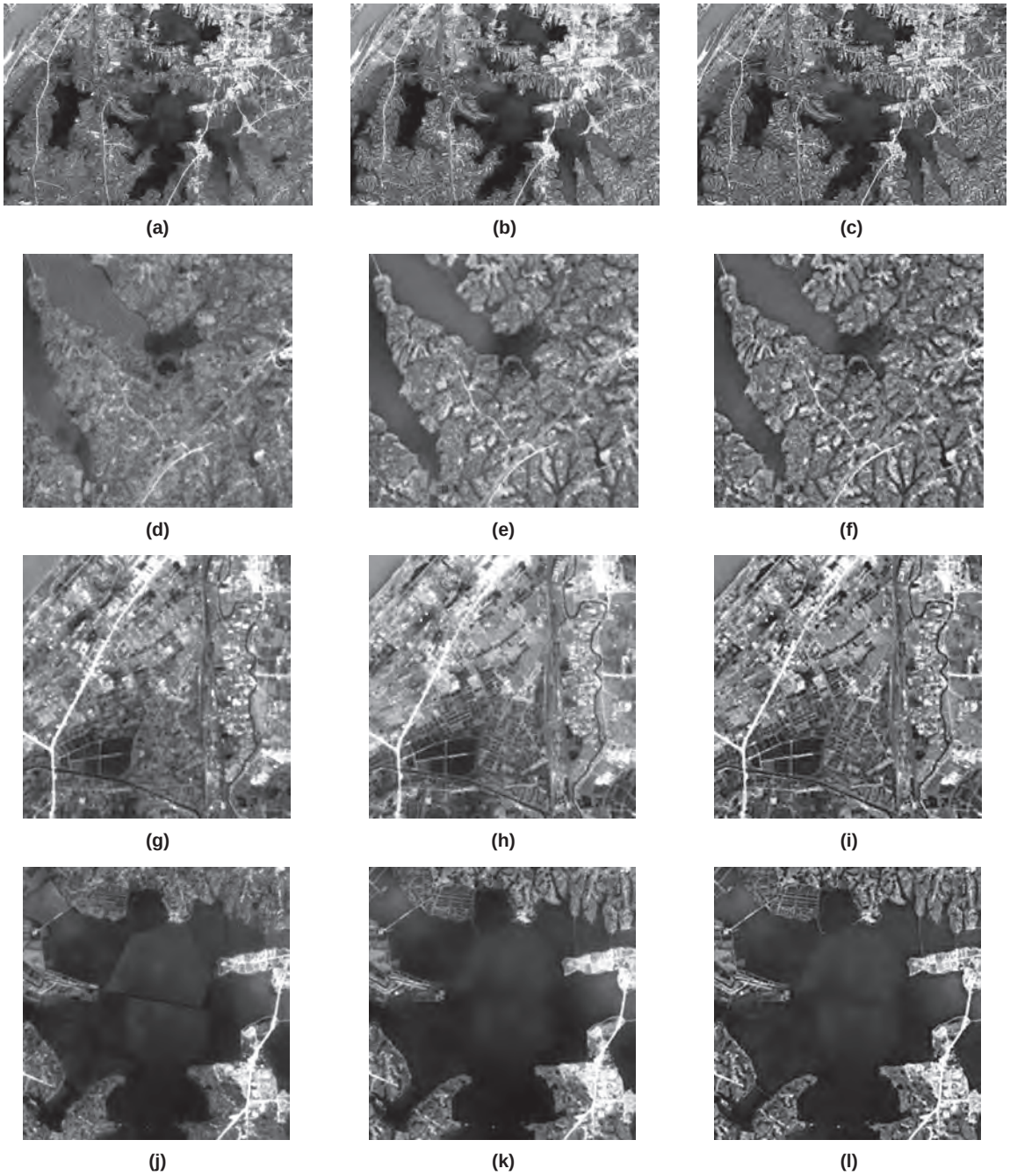

Figure 3. Results of the first clear image mainly covered by forest and water: (a) The original image; (d), (g), and (j) are the three detailed regions cropped from (a); (b) The result of MGF, $\alpha_{MGF} = 0.2$; (e), (h), and (k) are the three detailed regions cropped from (b); (c) The result of CMGF, $\alpha_{CMGF} = 0.2$; and (f), (i), and (l) are the three detailed regions cropped from (c).

Table 1. Means and Variances of Different Channels

		Band 1	Band 2	Band 3
Mean	Original	93.853	75.631	72.860
	MGF	93.854	75.629	72.859
	CMGF	93.852	75.630	72.860
Variance	Original	5.603	7.712	14.163
	MGF	5.227	7.269	13.371
	CMGF	6.610	7.898	14.222

However, the variance increases in the result of CMGF, especially in band 1, and the shorter the wavelength, the larger the increment. This outcome indicates that the results of MGF are not globally optimal; the MHP constraint works for obtaining the global optimum because it regularizes the intensity adjustment in each iteration, and then the fluctuation of the result data range is limited.

In Figure 4, the ground is almost covered by farmland, ponds and vegetation. The edges of farmland are all enhanced in the results of both MGF and CMGF, as shown in

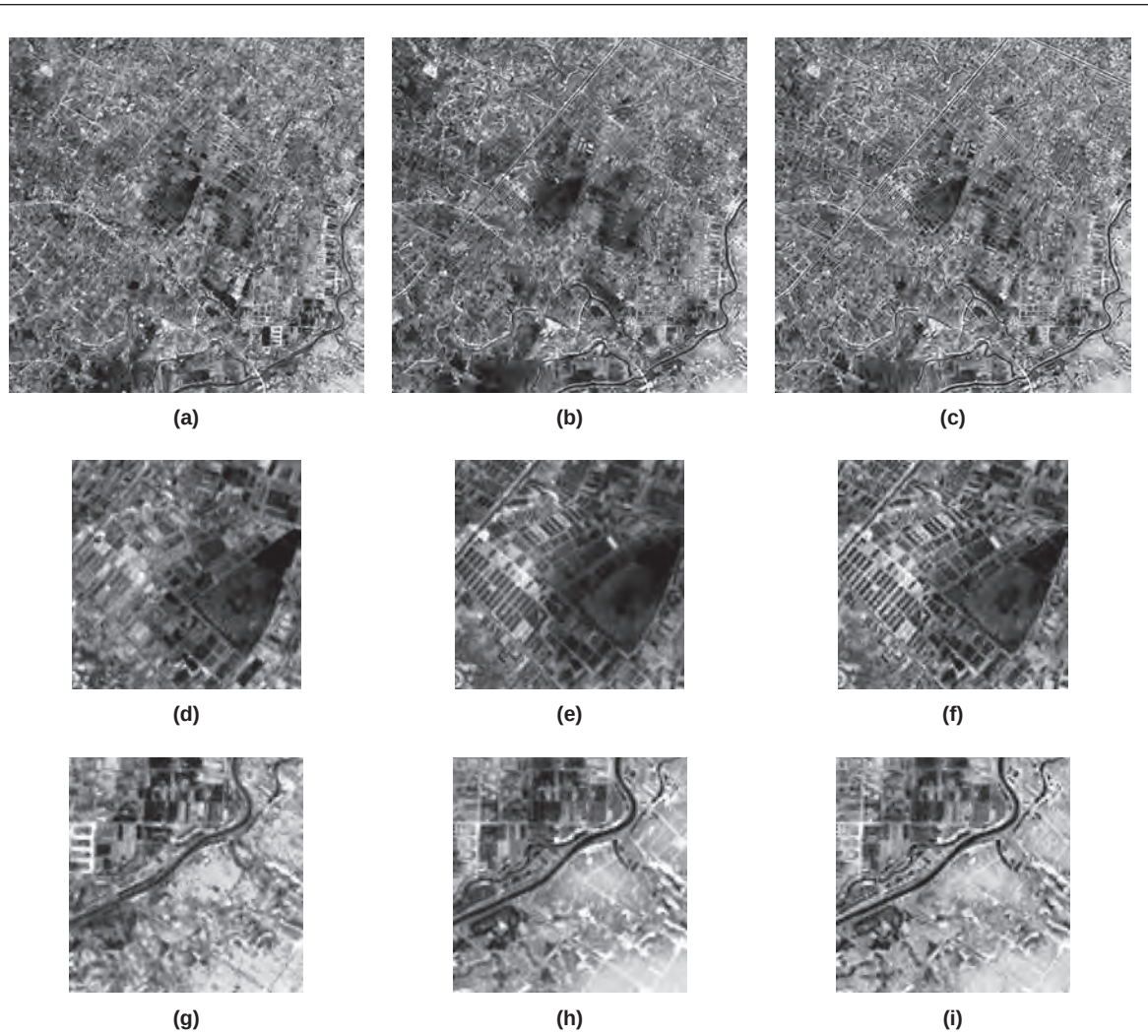

Figure 4. Results of the third clear scene mainly covered by farmland: (a) The original image; (d) and (g) are the two detailed regions cropped from (a); (b) The result of MGF, $\alpha_{MGF} = 0.2$; (e) and (h) are the two detailed regions cropped from (b); (c) The result of CMGF, $\alpha_{CMGF} = 0.2$; and (f) and (i) are the two detailed regions cropped from (c).

Figure 4b, 4c, 4e, and 4f. Since wetlands and water areas are both dark in band 7, the contrast in the original visible bands is attenuated in the fused result of MGF. Therefore, the edges of some areas with high humidity, such as ponds and vegetation areas, are smoothed in Figure 4b, 4e, and 4h. The results of CMGF retain the clear edges of humid areas and suppress noise in fields, as shown in Figure 4c, 4f, and 4i.

In order to assess the ability of GF methods to enhance image gradient quantitatively, we took an image structure measurement, Metric Q, as the evaluation factor. The definition of Metric Q is expressed as:

$$Q = s_1 \frac{s_1 - s_2}{s_1 + s_2} \tag{19}$$

where, s_1 and s_2 are the singular values of the gradient matrix $\mathbf{G}$ over an $N \times N$ window w_i in an image, and s_1 is intimately related to the sharpness of the local region. $\mathbf{G}$ is defined as:

$$\mathbf{G} = \begin{bmatrix} \vdots & \vdots \\ g_x(k) & g_y(k) \\ \vdots & \vdots \end{bmatrix}, k \in w_i \tag{20}$$

where k denotes the k^{th} pixel in the window w_i. Metric Q is properly correlated with the noise level, sharpness, and intensity contrast of the structured regions of an image without any prior knowledge (Zhu *et al.*, 2010). It reveals the signal-to-noise ratio of an image, so the larger the Metric Q, the better the result. Table 2 lists the Metric Q values of the two clear data sets and the GF results of them. It can be seen that the Metric Q values of the GF results are greatly increased compared to the original values. Results of MGF and CMGF are almost equal, and CMGF performs a little better than MGF. With the exception of band 1 of the clear data 1, a higher value appears in the result of MGF, as smoothing in water areas raises the performance of Metric Q.

Hazy Data

After investigating the application of gradient-based fusion models on clear scenes, we then investigated the effect of these models when applied to hazy scenes. Three images partially covered by haze and thin cloud were selected as shown in Plate 1, Figure 5, and Figure 6. The main ground types covered by these scenes were urban areas, lakes, rivers, farmland, and bare soil.

From the overall scenes and detailed regions in Plate 1, we can see that edges and detailed textures are both

TABLE 2. METRIC Q OF THE CLEAR DATA

		Band 1	Band 2	Band 3
Clear data 1 (water)	Original	0.0141	0.0217	0.0304
	MGF	0.0300	0.0312	0.0321
	CMGF	0.0286	0.0316	0.0350
Clear data 2 (farmland)	Original	0.0211	0.0245	0.0292
	MGF	0.0257	0.0258	0.0256
	CMGF	0.0267	0.0270	0.0370

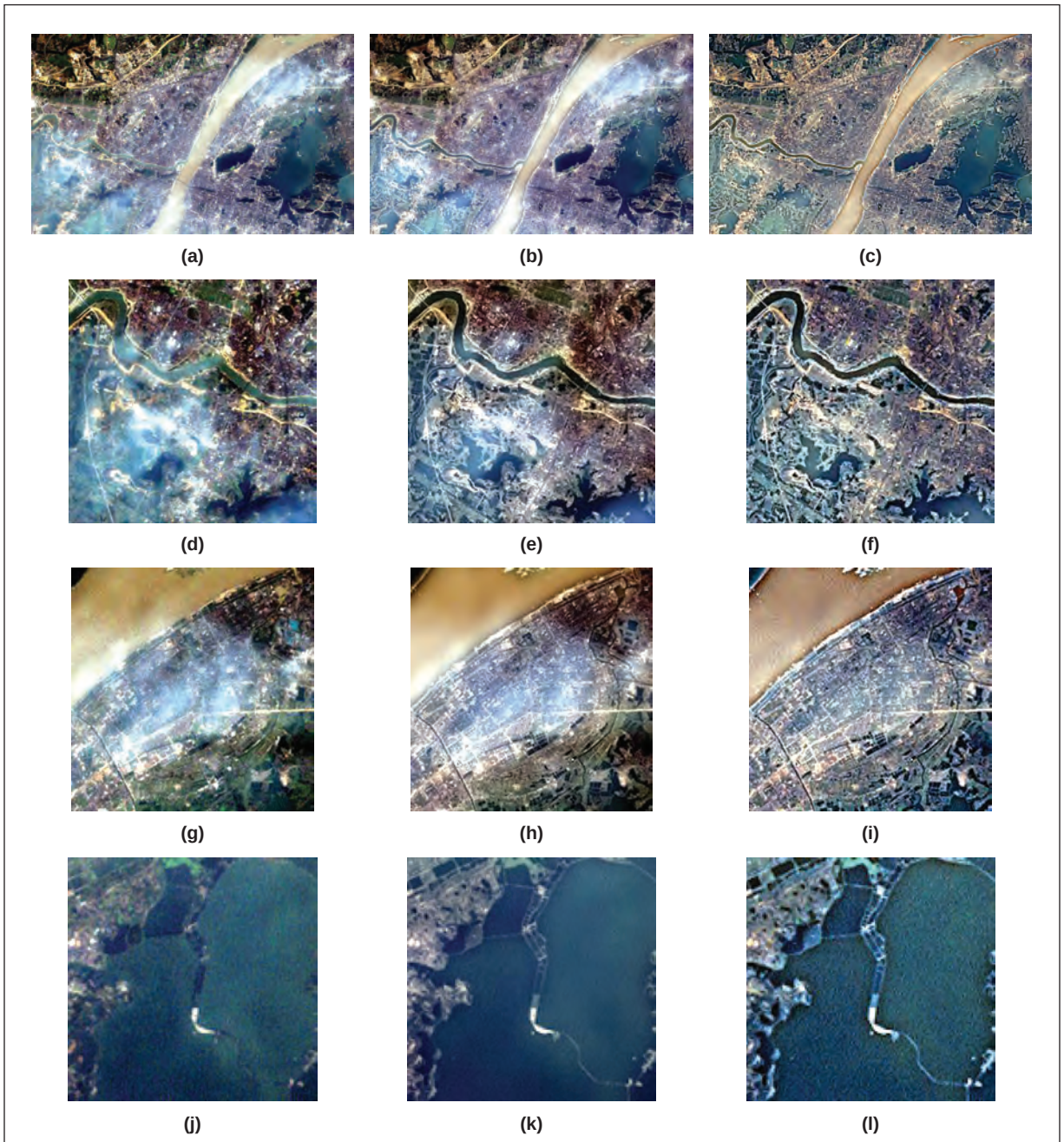

Plate 1. Results of the first hazy scene mainly covered by urban areas: (a) The original image; (d), (g), and (j) are the three detailed regions cropped from (a); (b) The result of MGF, $\alpha_{MGF} = 0.01$; (e), (h) and (k) are the three detailed regions cropped from (b); and (c) The result of CMGF, $\alpha_{CMGF} = 0.01$; (f), (i) and (l) are the three detailed regions cropped from (c).

enhanced in the results of MGF and CMGF, no matter what the land surface types. In general, the CMGF outperforms MGF in hazy regions, as the three detailed regions cropped from the overall scenes show in Plate 1. This result is because the MHP constraint in CMGF takes the band correlation into consideration. Gradients are properly enhanced while avoiding spectral distortion. However, the MHP constraint also enhances the tiny intensity contrast in the isotropic areas, as well as enhancing edges and textures. As the third detailed region in Plate 1l shows, even though the shape of the bridge on the water surface is highlighted, the heterogeneity of the water is also enhanced, which brings some noise into the result. With the reduction of the parameter, the contrast of the whole image grows, which means that the haze contamination is attenuated while the homogeneity is depressed. Therefore, the parameter α is the key to balance the contrast and the homogeneity. In the experiment, α was determined by considering the visual quality and the Metric Q of the image.

The results of CMGF in Figure 5 and Figure 6 are more satisfactory for haze and thin cloud removal than the results of MGF. All the edges and textures are clear enough to be distinguished and the overall gray tone is preserved. In particular, the river shows salient changes in the shape and tone in both Figure 5 and Figure 6. Note that the areas immediately adjacent to the river are bright, and the center of the river is dark in the fused images. Thus, the river in the fused results seems narrow when compared with the original scene. This result is due to the water quality of the Yangtze River. Water in the Yangtze River is rich in suspended silt, so that the color of the river itself resembles the color of bare soil. The components of the areas immediately adjacent to the river are complicated and highly affected by human activities. Thus, the corresponding intensity of these

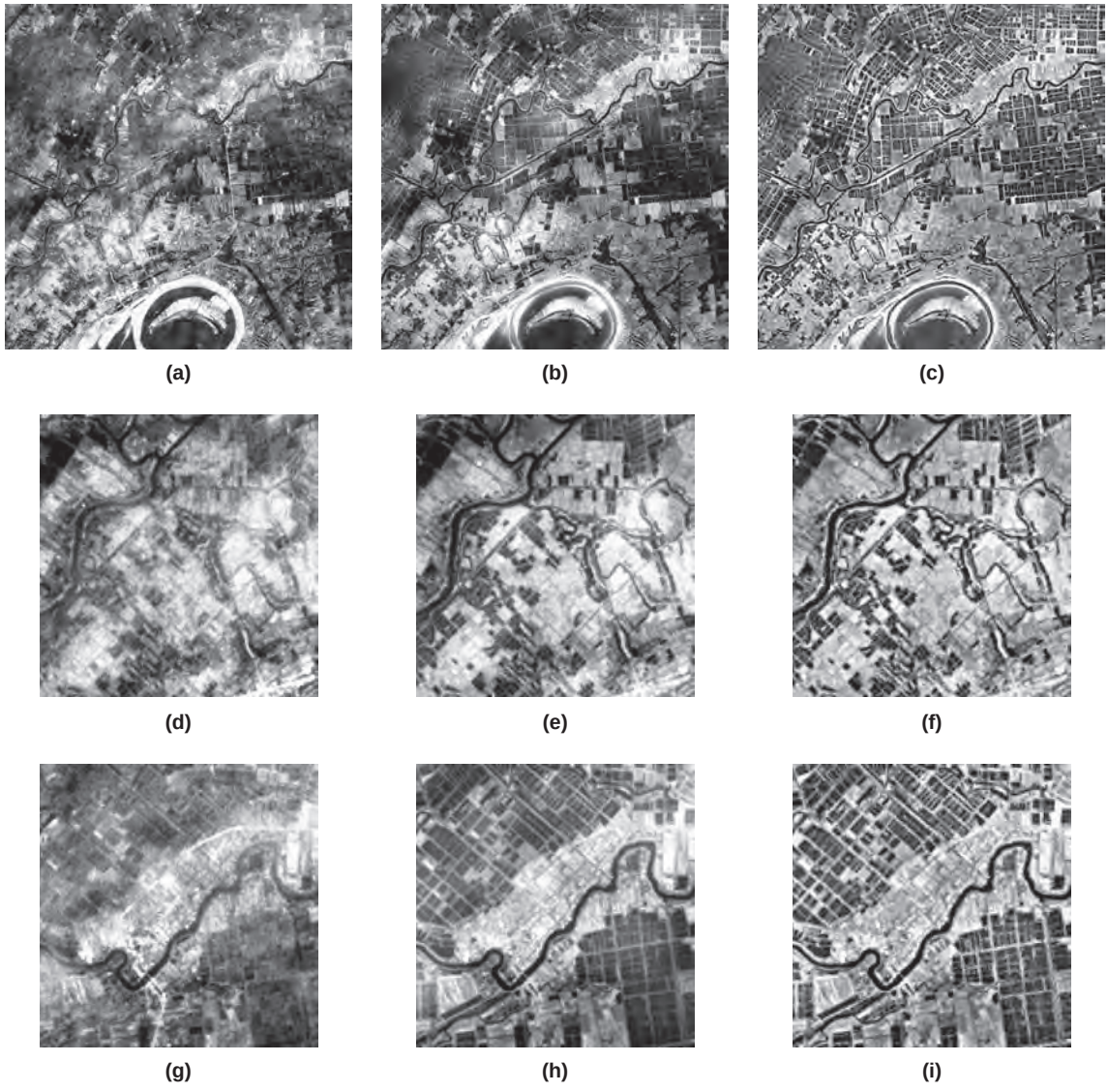

Figure 5. Results of the second hazy scene mainly covered by farmland and bare soil: (a) The original image; (d) and (g) are the two detailed regions cropped from (a); (b) The result of MGF, $\alpha_{MGF} = 0.01$; (e) and (h) are the two detailed regions cropped from (b); (c) The result of CMGF, $\alpha_{CMGF} = 0.01$; (f) and (i) are the three detailed regions cropped from (c).

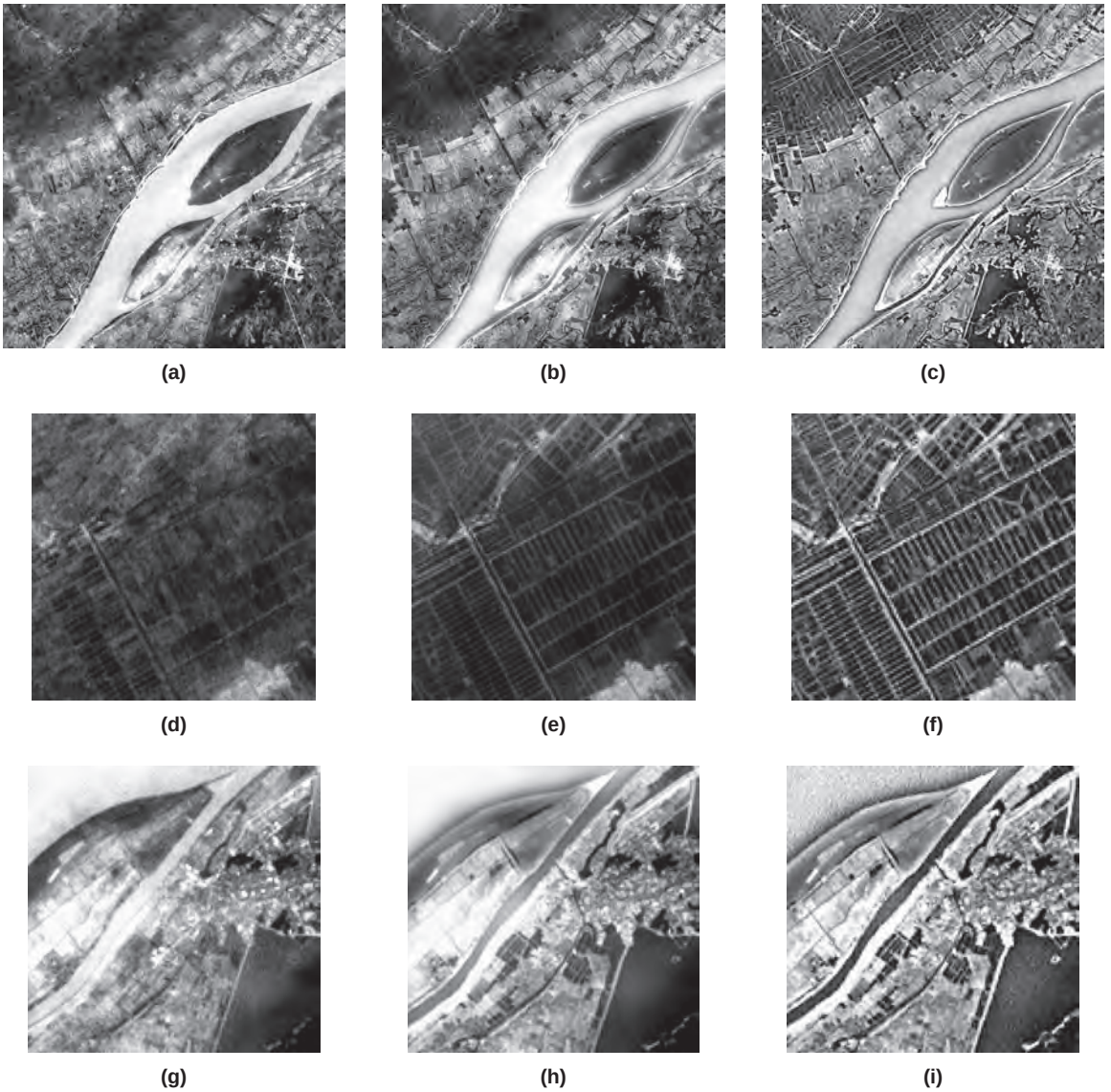

Figure 6. Results of the third hazy scene mainly covered by farmland, bare soil and the Yangtze River: (a) The original image; (d) and (g) are the two detailed regions cropped from (a); (b) The result of MGF, $\alpha_{MGF} = 0.01$; (e) and (h) are the two detailed regions cropped from (b); (c) The result of CMGF, $\alpha_{CMGF} = 0.01$; and (f) and (i) are the two detailed regions cropped from (c).

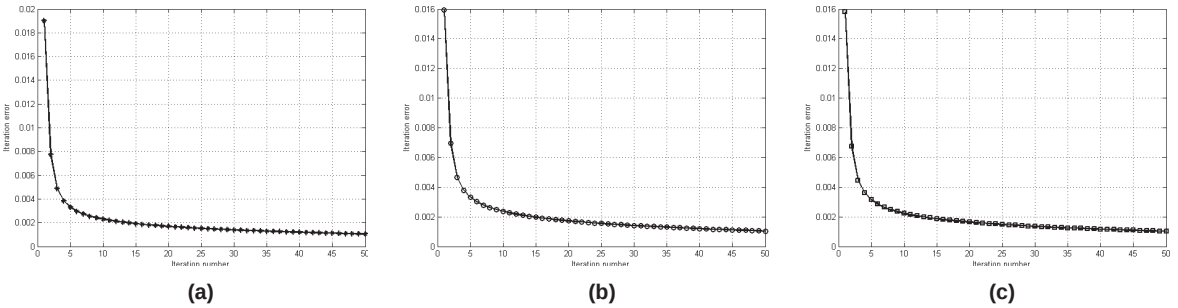

Figure 7. Iteration error curves of the above three hazy data.

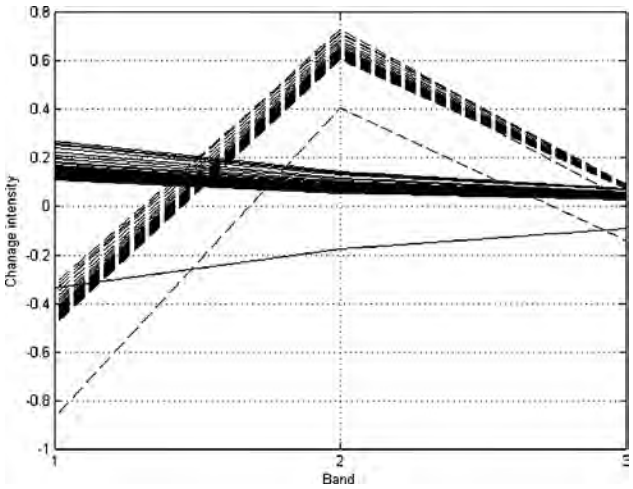

Figure 8. The intensity change curves of 50 iterations: the dashed curves represent the intensity change calculated from the variational function; the solid curves represent the adjusted intensity change after adding the MHP constraint.

areas is relatively high, approaching the intensity of man-made urban areas, so that vast contrast appears in the areas immediately adjacent to the river. This reason also explains the halo appearing in Figure 6b, 6h, 6c, and 6i. Moreover, it is noted that the change in the appearance of the Yangtze River is more distinct in the result of CMGF than the result of MGF. This result can also be attributed to the MHP constraint, which attenuates the atmospheric effect over the Yangtze River.

Metric Q was also employed to quantitatively assess the GF results of the hazy data, as listed in Table 3. A large increase in Metric Q appears in all the bands of the hazy image, and the CMGF outperforms the MGF, especially in the first two bands. In the result of CMGF, it is worth noting that the increase in Metric Q becomes less as the wavelength increases. This change agrees with the changing tendency of atmospheric effects between bands.

Convergence of CMGF

As a forcible constraint was added in the iteration of the CMGF solving process, we need to discuss the convergence of CMGF. The relationship curves of the iteration error and the iteration number are shown in Figure 7a, 7b, and 7c, corresponding to the above three hazy data sets. No matter what the land-cover types, the CMGF converges well within 50 iterations. This process suggests that CMGF is robust and does not take too much time to reach a stable solution.

Furthermore, in order to observe the impact of the MHP constraint on the intensity change quantitatively, we randomly selected a pixel from the first hazy data and tracked its intensity change in the three visible bands, as shown in Figure 8. The dashed curves represent the intensity change calculated in the variational equation solving process; they are totally composed of mathematical values. The solid curves show the adjusted intensity change after adding the MHP constraint. It can be seen that in each iteration the intensity change in the visible bands agrees with the negative exponent relation. This result means that the intensity changes in accordance with the physical characteristics of the atmospheric scattering. Therefore, stable and haze-free results can be obtained by the CMGF method.

TABLE 3. METRIC Q OF THE HAZY DATA

		Band 1	Band 2	Band 3
Hazy data 1 (urban)	Original	0.0117	0.0166	0.0202
	MGF	0.0232	0.0268	0.0291
	CMGF	0.0357	0.0282	0.0262
Hazy data 2 (farmland and bare soil)	Original	0.0182	0.0265	0.0396
	MGF	0.0451	0.0494	0.0582
	CMGF	0.0666	0.0667	0.0520
Hazy data 3 (farmland and Yangtze River)	Original	0.0114	0.0187	0.0344
	MGF	0.0585	0.0604	0.0690
	CMGF	0.0611	0.0685	0.0627

Conclusions

The gradient-based fusion of visible and SWIR imagery is an effective method for the enhancement of spatial information and restoration of true color. The MGF is a practical model of GF, designed for single-band imagery; however, it fails to de-haze hazy images. In order to implement dehazing, the MHP constraint based on the band correlation was introduced. With the inclusion of the MHP constraint, the CMGF is not only able to enhance the gradient, such as the MGF, but is also able to de-haze the hazy image. Since spatial and spectral information are integrated into one image, the results are endued with true color and clear edges, without haze contamination.

Both clear and hazy Landsat ETM+ data were used in the tests. The qualitative and quantitative experimental results demonstrated that the fused image of the CMGF is better than that of the MGF. Besides, it was noted that an appropriate parameter α, balancing the contrast and the homogeneity in the variational function, is important for yielding a good result. The smaller α is, the sharper the gradients of the fused image are, and the more homogeneity is destroyed. How to automatically select a proper parameter α will be the subject of our future work. Furthermore, technically, the proposed method has no limitations with regard to image resolution. Experiments on remote sensing data with other resolutions will be implemented in the future.

References

Aksoy, S., K. Koperski, C. Tusk, and G. Marchisio, 2009. Land cover classification withmulti-sensor fusion of partly missing data, *Photogrammetric Engineering & Remote Sensing*, 75(5): \ 577–593.

Ballester, C., V. Caselles, L. Igual, J. Verdera, and B. Rouge, 2006. A variational model for P+XS image fusion, *International Journal of Computer Vision*, 69(1):43–58.

Bucholtz, A., 1995. Rayleigh-scattering calculations for the terrestrial atmosphere, *Applied Optics*, 34(15):2765–2773.

Burt, P.J., and R.J. Kolczynski, 1993. Enhanced image capture through fusion, *Proceedings of the Fourth International Conference on \ Computer Vision*, Berlin, Germany, pp. 173–182.

Chavez, P.S., 1996. Image-based atmospheric corrections-revisited and improved, *Photogrammetric Engineering & Remote Sensing*, 62(9):1025–1035.

Crist, E.P., and R.C. Cicone, 1984. A physically-based transformation of Thematic Mapper data-the TM tasseled cap, *IEEE Transactions on Geoscience and Remote Sensing*, GE-22(3):256–263.

Davis, J.W., and V. Sharma, 2007. Background-subtraction using contour-based fusion of thermal and visible imagery, *Computer Vision and Image Understanding*, 106(2-3):162–182.

Davranche, A., G. Lefebvre and B. Poulin, 2009. Radiometric normalization of SPOT-5 scenes: 6S atmospheric model versus pseudo-invariant features, *Photogrammetric Engineering & Remote Sensing*, 75(6):723–728.

Fallah-Adl, H., J. Jájá, and S. Liang, 1997. Fast algorithms for estimating aerosol optical depth and correcting thematic mapper imagery, *The Journal of Supercomputing*, 10(4):315–329.

Fallah-Adl, H., J. Jájá, S. Liang, J. Townshend, and Y.J.Kaufman, 1996. Fast algorithms for removing atmospheric effects from satellite images, *IEEE Computational Science & Engineering*, 3(2):66–77.

Gadallah, F.L., F. Csillag, and E.J.M. Smith, 2000. Destriping multisensor imagery with moment matching, *International Journal of Remote Sensing*, 21(12):2505–2511.

He, X.Y., J.B. Hu, W. Chen, and X.Y. Li, 2010. Haze removal based on advanced haze-optimized transformation (AHOT) for multispectral imagery, *International Journal of Remote Sensing*, 31(20):5331–5348.

Huang, X., L. Zhang and P. Li, 2007. Classification and extraction of spatial features in urban areas using high resolution multispectral imagery, *IEEE Geoscience and Remote Sensing Letters*, 4(2): 260–264.

Huang, X., and L. Zhang, 2008. An adaptive mean-shift analysis approach for object extraction and classification from urban hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 46(12):4173–4185.

Jacobsen, A., K.B. Heidebrecht, and A.F.H. Goetz, 2000. Assessing the quality of the radiometric and spectral calibration of casi data and retrieval of surface reflectance factors, *Photogrammetric Engineering & Remote Sensing*, 66(9): 1083–1092.

Karnieli, A., Y.J. Kaufman, L. Remer, and A. Wald, 2001. AFRI - Aerosol free vegetation index, *Remote Sensing of Environment*, 77(1):10–21.

Kaufman, Y.J., 1994. Detection of forests using mid-IR reflectance, An application for aerosol studies, *IEEE Transactions on Geoscience and Remote Sensing*, 32(3):672–683.

Kaufman, Y.J., A.E. Wald, L.A. Remer, B. Gao, R. Li, and L. Flynn, 1997. The MODIS 2.1-μm channel-correlation with visible reflectance for use in remote sensing of aerosol, *IEEE Transactions on Geoscience and Remote Sensing*, 35(5):1286–1298.

Liang, S., 2001. Atmospheric correction of Landsat ETM+ land surface imagery-Part I: Methods, *IEEE Transactions on Geoscience and Remote Sensing*, 39(11):2490–2498.

Mahiny, A.S., and B.J. Turner, 2007. A comparison of four common atmospheric correction methods, *Photogrammetric Engineering & Remote Sensing*, 73(4):361–368.

McCartney, E.J., 1976. *Optics of the Atmosphere: Scattering by Molecules and Particles*, New York, John Wiley and Sons, Inc., 421 p.

Nayar, S.K., and S.G. Narasimhan, 1999. Vision in bad weather, *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, Kerkyra, Greece, pp. 820–827

Norjamaki, I., and T. Tokola, 2007. Comparison of atmospheric correction methods in mapping timber volume with multitemporal Landsat images in Kainuu, Finland, *Photogrammetric Engineering & Remote Sensing*, 73(2):155–163.

Petrovic, V.S., and C.S. Xydeas, 2004. Gradient-based multiresolution image fusion, *IEEE Transactions on Image Processing*, 13(2):228–237.

Piella, G., 2009. Image fusion for enhanced visualization: a variational approach, *International Journal of Computer Vision*, 83(1):1–11.

Pohl, C., and J. Van Genderen, 1998. Review article: Multisensor image fusion in remote sensing: Concepts, methods and applications, *International Journal of Remote Sensing*, 19(5): 823–854.

Raskar, R., A. Ilie, and J. Yu, 2005. Image fusion for context enhancement and video surrealism, *Proceedings of SIGGRAPH '05 ACM SIGGRAPH 2005 Courses*, pp. 4-es.

Richter, R., 1996a. Atmospheric correction of satellite data with haze removal including a haze/clear transition region, *Computers & Geosciences*, 22(6):675–681

Richter, R., 1996b. A spatially adaptive fast atmospheric correction algorithm, *International Journal of Remote Sensing*, 17(6): 1201–1214.

Socolinsky, D.A., and L.B. Wolff, 1999. A new visualization paradigm for multispectral imagery and data fusion, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Fort Collins, Colorado.

Socolinsky, D.A., and L.B. Wolff, 2002. Multispectral image visualization through first-order fusion, *IEEE Transactions on Image Processing*, 11(8):923–931.

Stark, J.A., 2000. Adaptive image contrast enhancement using generalizations of histogram equalization, *IEEE Transactions on Image Processing*, 9(5):889–896.

Su, H., D. Karna, E. Fraim, M. Fitzgerald, R. Dominguez, J.S. Myers, B. Coffland, L.R. Handley and T. Mace, 2006. Evaluation of eelgrass beds mapping using a high-resolution airborne multispectral scanner, *Photogrammetric Engineering & Remote Sensing*, 72(7):789–797.

Wu, J., D. Wang, and M.E. Bauer, 2005. Image-based atmospheric correction of QuickBird imagery of Minnesota cropland, *Remote Sensing of Environment*, 99(3):315–325.

Zhang, L., B. Du, and Y. Zhong, 2010. Hybrid detectors based on selective endmembers, *IEEE Transactions on Geoscience and Remote Sensing*, 48(6):2633–2646.

Zhang, Y., and B. Guindon, 2003. Quantitative assessment of a haze suppression methodology for satellite imagery: Effect on land cover classification performance, *IEEE Transactions on Geoscience and Remote Sensing*, 41(5):1082–1089.

Zhang, Y., B. Guindon, and J. Cihlar, 2002. An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images, *Remote Sensing of Environment*, 82(2-3):173–187.

Zhou, Z., Y. Li, H. Shi, N. Ma, C. He, and P. Zhang, 2010. A fast variational fusion approach for pan-sharpening, *Proceedings of the 2010 IEEE 10th International Conference on Signal Processing*, Beijing, pp. 1110–1113.

Zhu, X., and P. Milanfar, 2010. Automatic parameter selection for denoising algorithms using a no-reference measure of image content, *IEEE Transactions on Image Processing*, 19(12): 3116–3132.

(Received 05 September 2011; accepted 12 December 2011; final version 26 January 2012)