Geocarto International
Vol. 23, No. 2, April 2008, 155–165

ISVR: an improved synthetic variable ratio method for image fusion

L. WANG†, X. CAO‡ and J. CHEN*‡
†Department of Geography, The State University of New York at Buffalo, USA
‡Key Laboratory of Environmental Change and Natural Disaster, Beijing Normal University, China

(*Received 4 July 2006; in final form 2 July 2007*)

An Improved Synthetic Variable Ratio (ISVR) fusion method is proposed to merge high spatial resolution panchromatic (Pan) images and multispectral (MS) images based on a simulation of the panchromatic image from the multispectral bands. Compared to the existing SVR (Synthetic Variable Ratio) family methods, the ISVR method manifests two major improvements: a simplified and physically meaningful scheme to derive the parameters necessary as required by SVR, and less computing power. Two sets of IKONOS Pan and MS images: one in urban area and another one in a forest area, were used to evaluate the effectiveness of classification-oriented ISVR method in comparison to the Principal Component Substitution (PCS), Synthetic Variable Ratio (SVR) and Gram-Schmidt Spectral Sharpening (GS) methods that are available in the ENVI software package. Results indicate the ISVR method achieves the best spectral fidelity to facilitate classification compared to PCS, SVR, and GS methods.

Keywords: Image fusion; synthetic variable ratio

1. Introduction

Image fusion is a technique aimed at producing new multispectral (MS) images that can inherit both the higher-spatial information from panchromatic (Pan) images and the higher spectral information embedded in the original lower-spatial resolution MS images. The technique emerged in mid-1980s and has effectively aided the development of a variety of techniques and applications, such as image sharpening, registration refining, stereo image generation, feature enhancement, image classification, change detection, and image visualization (Pohl and van Genderen 1998, Zhang 2004). Traditional image fusion methods can be grouped in three categories: (1) transformation-based methods, such as Intensity, Hue, Saturation (IHS) and Principal Component Substitution (PCS); (2) arithmetic combination based methods such as the Brovey Transform, SVR, and Ratio Enhancement (RE); and (3) filter-based methods including High-Pass Filer (HPF) (Chavez *et al.* 1991), Smoothing Filter-based Intensity Modulation (SFIM) and wavelet-based methods.

*Corresponding author. Email: chen.jin@nies.go.jp

Geocarto International
ISSN 1010-6049 print/ISSN 1752-0762 online © 2008 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/10106040701204198

A comprehensive review of different methods can be found in Pohl and van Genderen (1998). Recently, several new methods and models were proposed. For example, context-driven method (Aiazzi *et al*. 2002), based on wavelet transform and Laplacian pyramid, which were all generalized into ARSIS concept (from French acronym: '*Amélioration de la Résolution Spatiale par Injection de Structures*', meaning 'spatial resolution enhancement by injection of structures') (Ranchin and Wald 2000; Ranchin *et al*. 2003).

Among these methods, the SVR method that was initially proposed by Munechika *et al*. (1993) and later modified by Zhang (1999) has presented some advantages over other methods in the process of merging higher spectral MS images and higher spatial panchromatic images. In addition, SVR requires less computing power than the wavelet transform does. The method has been successfully applied to detection of urban housing development (Zhang 2001). Nevertheless, there are several restrictions with the SVR method. Among them, one major obstacle with the SVR method lies in the difficulty to determine an optimal set of parameters as is imperative for generating simulated panchromatic images. Common practice for deriving such parameters is through a multiple regression analysis, which is usually time consuming and inconsistent as the input image changes. In this study, in order to circumvent this problem, an ISVR method was proposed with an aim to provide a simplified manner to derive universal parameters that are invariant to all the images coming from the same data source. As an evaluation, the ISVR method was compared with PCS, Zhang's (1999) SVR and Gram-Schmidt GS method is available in the ENVI 4.1 software package.

2. ISVR method for image fusion

2.1 *SVR*

Munechika *et al*. (1993) developed an SVR method to merge an MS image and a high resolution panchromatic image as described in the following equation:

$$XSP_i = Pan_{\mathrm{H}} \frac{XS_{\mathrm{L}i}}{Pan_{\mathrm{LSyn}}}, \tag{1}$$

where XSP_i is the grey value of the ith band of the merged image, Pan_{H} is the grey value of the original high spatial resolution image, $XS_{\mathrm{L}i}$ is the grey value of the ith band of the original MS image, and Pan_{LSyn} is the grey value of the low resolution synthetically panchromatic image by the following simulation equation proposed by Suits *et al*. (1988):

$$Pan_{\mathrm{LSyn}} = \sum_{i=1}^{4} \varphi_i XS_{\mathrm{L}i}. \tag{2}$$

In the study reported by Munechika *et al*. (1993), TM Bands 1, 2, 3 and 4 were used as $XS_{\mathrm{L}i}$ to calculate Pan_{LSyn}. In order to obtain parameters φ_i, Munechika *et al*. (1993) adopted a modified atmospheric model. The parameters φ_i were then calculated through a regression analysis between the values simulated through the atmospheric model and then measured for five typical land cover types: urban, soil, water, trees and grass. After construction of the Pan_{LSyn}, a linear histogram match was used to force the original SPOT Pan image to match the Pan_{LSyn} in order to eliminate atmospheric and illumination differences.

In order to derive a more stable φ_i, Zhang (1999) modified the SVR method as follows:

$$XSP_i = Pan_{\mathrm{H}} \frac{XS_{\mathrm{H}i}}{Pan_{\mathrm{HSyn}}}, \tag{3}$$

where XSP_i and Pan_{H} remains the same as in equation (1), while $XS_{\mathrm{H}i}$ is the grey value of the *i*th band of the TM image resampled to the same pixel size as the original SPOT Pan image by bilinear interpolation, and Pan_{HSyn} is the grey value of the high resolution synthetically panchromatic image simulated through the following equation:

$$Pan_{\mathrm{HSyn}} = \sum \varphi_i XS_{\mathrm{H}i}. \tag{4}$$

In order to calculate the parameters φ_i, the following equation was used:

$$Pan_{\mathrm{H}} = \sum \varphi_i XS_{\mathrm{H}i}. \tag{5}$$

In this equation, the parameters φ_i were calculated directly through multiple regression analysis of the original SPOT Pan image Pan_{H} and TM MS bands $XS_{\mathrm{H}i}$ instead of only from the five land cover types as is the case with Munechika *et al.* (1993).

According to Zhang's (1999) SVR modification, the parameters φ_i are mainly determined by the multiple regression analysis. The poorer the correlation between Pan_{H} and the MS bands $XS_{\mathrm{H}i}$ is, the less spectral contents in the MS bands $XS_{\mathrm{H}i}$ will be integrated in the calculation of Pan_{HSyn}. However, some restrictions still lie with the modified SVR method. First, the correlation dependence tends to create colour distortion in the fused image in some cases. Second, whenever a new image is to be fused, the multiple regression analysis has to be rerun in order to derive appropriate values for the parameter φ_i. Thus, computation burden is added. Finally, the multiple regression analysis is still time consuming, particularly when fusion for a large image has to be carried out.

2.2 *ISVR*

In order to address the aforementioned limitations with SVR and the modified SVR, an ISVR method was proposed in this study in an attempt to obtain the parameters φ_i in a more simplified manner than with regression analysis. Our method draws upon a basic assumption: radiance of the synthesized panchromatic band can be obtained through integration of the radiance of MS bands, where the MS bands cover the same spectra range as the panchromatic band, and the spectra response function of the panchromatic band is identical to that of the MS bands. Based on this assumption, the radiance of the synthetically panchromatic image can be considered as the spectra integral in the spectra range of the original panchromatic band. If we decompose the integral to areas of each MS band as well as the gaps among the MS bands, we can assume that the radiance of each MS band is equal to the rectangular area while the radiance of the gaps approximately equals to the trapezoid areas between the rectangles (figure 1). Here, the area of trapezoids is dependent upon the neighbouring rectangles and their spectra range and areas can be calculated according to the equations (6), (7) and (8). After the radiance area of gaps was obtained, the integral (or the radiance of simulated panchromatic image) can be calculated by summing all the areas together.

The basic equation of ISVR is the same as Zhang's modification (equation (3)), whereas the Pan_{HSyn} is calculated through the following equations

$$\begin{aligned} Pan_{\mathrm{HSyn}} = {} & XS_{\mathrm{H1}} + XS_{\mathrm{H2}} + \cdots + XS_{\mathrm{Hn}} + \left(\frac{XS_{\mathrm{H1}}}{\lambda_{1,2} - \lambda_{1,1}} + \frac{XS_{\mathrm{H2}}}{\lambda_{2,2} - \lambda_{2,1}} \right) \\ & \times \frac{\lambda_{2,1} - \lambda_{1,2}}{2} + \left(\frac{XS_{\mathrm{H2}}}{\lambda_{2,2} - \lambda_{2,1}} + \frac{XS_{\mathrm{H3}}}{\lambda_{3,2} - \lambda_{3,1}} \right) \times \frac{\lambda_{3,1} - \lambda_{2,2}}{2} + \cdots \\ & + \left(\frac{XS_{\mathrm{H(n-1)}}}{\lambda_{(\mathrm{n}-1),2} - \lambda_{(\mathrm{n}-1),1}} + \frac{XS_{\mathrm{Hn}}}{\lambda_{\mathrm{n},2} - \lambda_{\mathrm{n},1}} \right) \times \frac{\lambda_{\mathrm{n},1} - \lambda_{(\mathrm{n}-1),2}}{2}, \end{aligned} \tag{6}$$

which can be also written as

$$Pan_{\mathrm{HSyn}} = \sum_{i=1}^{n} \varphi_i XS_{\mathrm{H}i}, \tag{7}$$

where

$$\begin{cases} \varphi_i = 1 + t_1 + t_2 \\ t_1 = \begin{cases} \frac{\lambda_{i,1} - \lambda_{(i-1),2}}{2(\lambda_{i,2} - \lambda_{i,1})} & (2 \le i < n) \\ 0 & (i = 1) \end{cases} \\ t_2 = \begin{cases} \frac{\lambda_{(i+1),1} - \lambda_{i,2}}{2(\lambda_{i,2} - \lambda_{i,1})} & (1 \le i < n) \\ 0 & (i = n) \end{cases} \end{cases}, \tag{8}$$

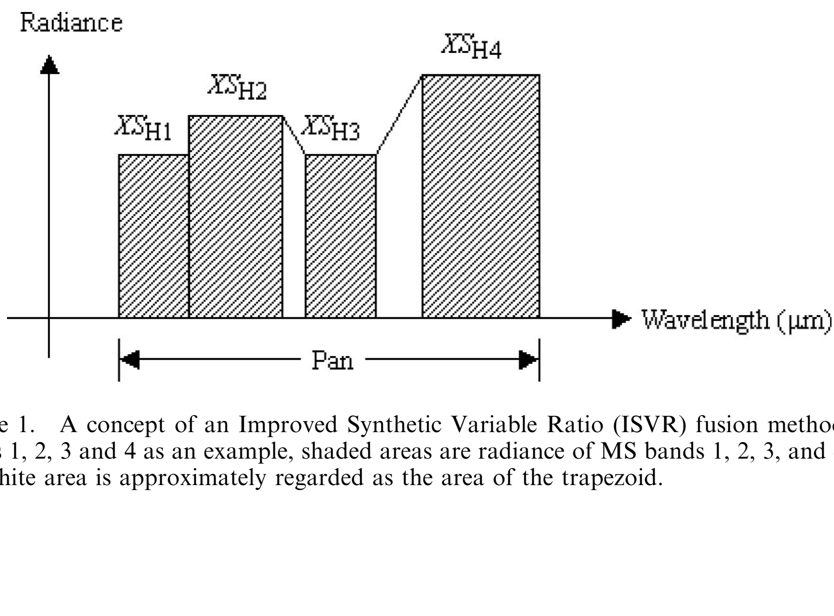

Figure 1. A concept of an Improved Synthetic Variable Ratio (ISVR) fusion method. MS bands 1, 2, 3 and 4 as an example, shaded areas are radiance of MS bands 1, 2, 3, and 4, and the white area is approximately regarded as the area of the trapezoid.

where $\lambda_{i,1}$ and $\lambda_{i,2}$ are two ends' wavelengths of *i*th band of original MS bands, and *n* is the total number of bands used to construct the panchromatic image.

As is shown, it is computationally convenient to calculate parameters φ_i when the wavelengths of each band involved in simulation of the panchromatic image are known. In addition, once the parameters φ_i are determined, they can be applied to other images coming from the same data source, e.g. IKONOS, Quickbird, TM etc. Thus, users do not need experience on a specific study area. Finally, the derived parameters have meaningful physical explanation.

2.3 *PCS and GS*

As a reference for examining the effectiveness of the proposed ISVR method, two popular methods in this category, PCS and Gram-Schmidt GS method, were adopted for the purposes of comparison. The first step of PCS was a traditional Principal Component (PC) transformation for the multispectral image. Then the high spatial resolution panchromatic band was used to replace the PC band 1 with a condition to preserve the original spectral range of the PC 1. An inverse PC transformation was then conducted with the replaced PC bands, resulting in a multispectral image with higher spatial resolution. With regards to the GS method, four steps were taken. A simulated panchromatic band was first generated using the multispectral bands. Then a Gram-Schmidt (GS) transformation was carried out for the stacked bands with the simulated panchromatic band as the first band and the remaining bands coming from the multispectral bands. Akin to the PCS method, the high spatial resolution panchromatic band was swapped with the first GS band, followed by an inverse GS transformation to derive the fused multispectral image at high spatial resolution.

3. Evaluation of image fusion methods

Wald *et al.* (1997) concluded that a good synthetic image must satisfy three criteria. First, where the synthetic image B_H^* is degraded to its original resolution L, it should largely resemble the original image B_L. Second, any synthetic image B_H^* should be as identical as possible to the image B_H that the corresponding sensor would observe with the highest spatial resolution H, if existent. Third, the MS set of synthetic images $B_H^*\hat{o}$ should be as identical as possible to the MS set of images B_H the corresponding sensor would observe with the highest spatial resolution H, if existent.

The three criteria lay out a basic scheme to assess various classification-oriented fusion methods. In previous studies, these criteria have been adapted to assess spectral quality for other fusion methods (Ranchin and Wald 2000, Ranchin *et al.* 2003). Since B_H is unavailable, Wald *et al.* (1997) suggested the following substitution method: the available images Pan A_H and MS B_L are degraded to A_L and B_S. For example, with regards to IKONOS images, the Pan and MS images would be degraded to 4 m and 16 m, and then the fusion method applied to A_L and B_S, which would produce a synthetic image B_L^*. This fused image B_L^* would then be compared to the original MS image B_L by means of the criteria. This comparison provides a way to assess the quality of B_L^*. It is assumed that this quality assessment is similar to that of the fused high-resolution image B_H^*.

4. Case study and results

In order to investigate the effectiveness of the proposed ISVR method, two sets of IKONOS Pan and MS images acquired over two representative landscapes were adopted for testing. The first set was in an urban setting (Beijing, China) where higher spatial resolution is always desired in various studies (figure 2(a) and (b)). Another was a mangrove forest image collected on the Caribbean coast of Panama as representative of the natural resource (figure 3(a) and (b)). In the pre-processing stage, IKONOS MS and Pan images were registered to the same coordinate system separately and the radiance was calculated from the digital numbers (DNs). Then IKONOS Pan and MS image were degraded to 4 m and 16 m respectively with a simple mathematic averaging algorithm. Finally, the degraded IKONOS MS image (16 m) was resampled by bilinear interpolation to 4 m for the follow-up ISVR

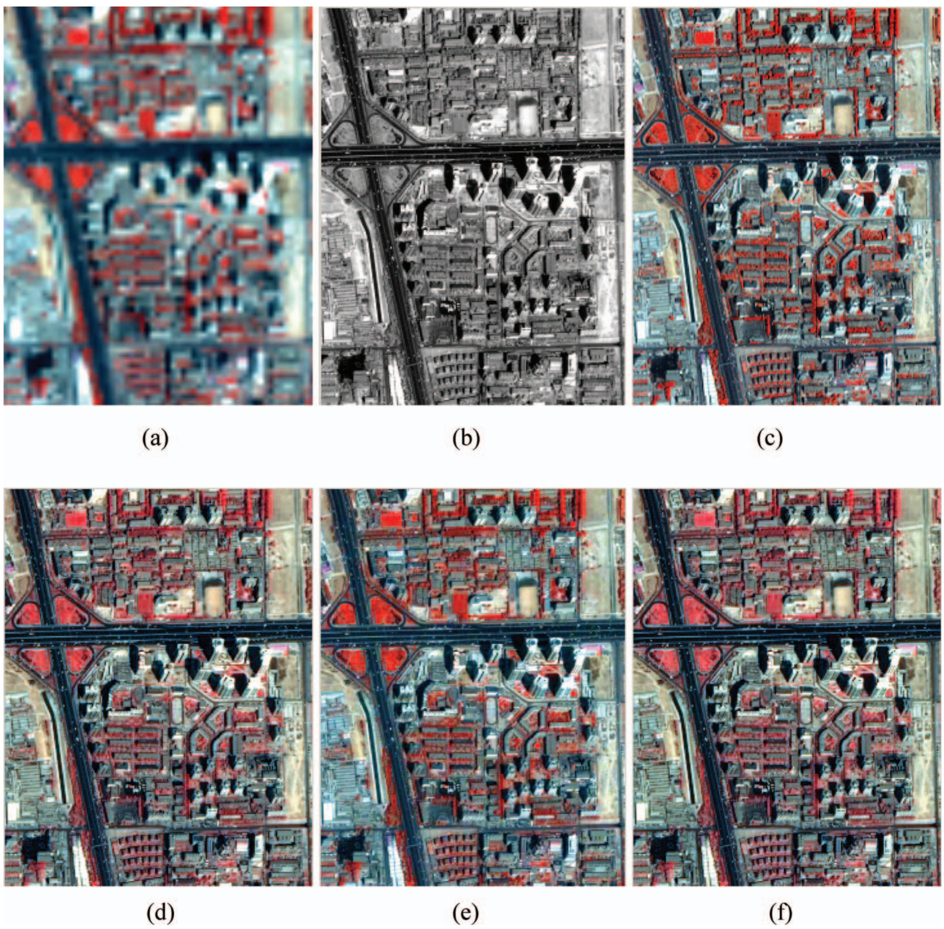

Figure 2. Comparison of the spectral and spatial effects of the fused images in Beijing, China. (a) Original IKONOS MS image with bands 4, 2, and 1 as RGB, (b) IKONOS Pan image, (c) the fused image by ISVR, (d) the fused image by PCS, (e) the fused image by Zhang's SVR, and (f) the fused image by GS method. Available in colour online.

calculation. To conduct an objective assessment, the same data were also fed to three other methods: PCS, Zhang's SVR and GS method, and histogram matching was applied to all fused images to preserve the mean values of the low resolution MS bands.

All the four bands of IKONOS MS were used to simulate Pan images from MS bands in order to be close to original Pan spectral range and avoid colour distortion (Zhang 2004). Given the spectra range of MS bands listed in table 1, we calculated parameters φ_i of the IKONOS MS bands, as listed in table 2. These values of φ_i can be treated as invariant to any IKONOS MS images. It should be noted that there is

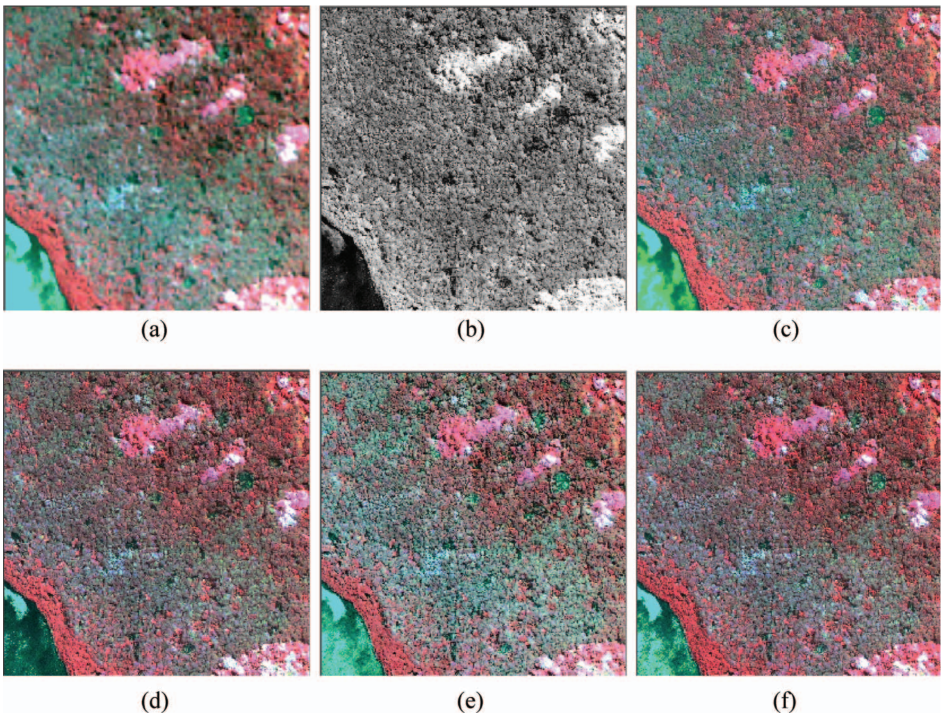

Figure 3. Comparison of the spectral and spatial effects of the fused images of mangrove forest. (a) Original IKONOS MS image with bands 4, 3, and 2 as RGB, (b) IKONOS Pan image, (c) the fused image by ISVR, (d) the fused image by PCS, (e) the fused image by Zhang's SVR, and (f) he fused image by GS method. Available in colour online.

Table 1. Parameters of IKONOS Pan and MS images.

Image	Band range (μm)	Spatial resolution
IKONOS MS image	Band 1: 0.445–0.516 Band 2: 0.506–0.595 Band 3: 0.632–0.698 Band 4: 0.757–0.853	4 m
IKONOS Pan image	0.45–0.90	1 m

an overlap between IKONOS band 1 and 2, so $\varphi_1 < 1$. Based on the derived parameters φ_i and equation (7) and (8), we created fused images as shown in figures 2(c) and 3(c), in which IKONOS MS bands 4, 2, 1 or 4, 3, 2 were used as the RGB input. In figures 2 and 3, we also presented the fused images by PCS, Zhang's SVR and GS method.

Tables 3 and 4 report some statistical indices on the relative discrepancies between fused images B_L^* and original images B_L for different methods. With regards to each band and each type of method, bias, standard deviation (*SD*) and root mean square error (RMSE) were calculated with the following equations and expressed in percentages that are relative to the mean radiance value of the original image B_L:

$$\begin{aligned} D_k &= B_{Lk}^* - B_{Lk} \\ Bias_k &= \frac{\overline{D}_k}{\overline{B}_{Lk}} \times 100 \\ SD_k &= \frac{\sqrt{\sum_{i=1}^{N} (D_k - \overline{D}_k)^2/(N-1)}}{\overline{B}_{Lk}} \times 100 \\ RMSE_k &= \frac{\sqrt{\overline{(D_k)^2}}}{\overline{B}_{Lk}} \times 100 \end{aligned} \tag{9}$$

Table 2. Parameters of IKONOS MS images for synthetic simulation of Pan images.

φ_1	φ_2	φ_3	φ_4
0.9296	1.1517	1.7273	1.3073

Table 3. Some statistics (percentages) and correlation coefficients between the synthetic and original images (image set 1: Beijing).

Band	Method	Bias	Standard deviation	RMSE	Correlation coefficient
Blue	ISVR	−0.19	3.2	3.2	0.87
	PCS	−0.16	4.6	4.6	0.82
	Zhang's SVR	−0.18	4.3	4.3	0.86
	GS	−0.11	3.7	3.7	0.81
Green	ISVR	−0.24	4.8	4.8	0.91
	PCS	−0.25	5.5	5.6	0.85
	Zhang's SVR	−0.27	5.3	5.3	0.86
	GS	−0.27	5.5	5.5	0.85
Red	ISVR	−0.33	7.5	7.5	0.91
	PCS	−0.40	7.3	7.3	0.86
	Zhang's SVR	−0.41	7.2	7.3	0.89
	GS	−0.44	8.6	8.6	0.85
NIR	ISVR	−0.43	8.8	8.8	0.92
	PCS	−0.45	8.6	8.6	0.87
	Zhang's SVR	−0.50	8.7	8.7	0.86
	GS	−0.46	9.5	9.5	0.84

Table 4. Some statistics (percentages) and correlation coefficients between the synthetic and original images (image set 2: mangrove forest).

Band	Method	Bias	Standard deviation	RMSE	Correlation coefficient
Blue	ISVR	−0.35	4.8	4.8	0.90
	PCS	−0.89	5.2	5.3	0.81
	Zhang's SVR	−0.64	5.4	6.0	0.90
	GS	−0.53	5.9	5.1	0.85
Green	ISVR	−0.42	5.1	5.1	0.88
	PCS	−0.77	6.4	6.4	0.85
	Zhang's SVR	−0.45	5.7	5.8	0.88
	GS	−0.69	5.5	5.5	0.83
Red	ISVR	−0.58	6.4	6.4	0.88
	PCS	−0.82	8.9	8.9	0.85
	Zhang's SVR	−0.76	7.8	7.5	0.84
	GS	−0.89	8.0	8.0	0.82
NIR	ISVR	−0.38	7.6	7.6	0.93
	PCS	−0.75	9.3	9.4	0.82
	Zhang's SVR	−0.54	8.5	8.6	0.89
	GS	−0.62	9.1	9.1	0.85

where D_k is the difference image of the kth band, B^*_{Lk} and B_{Lk} are the fused image and original MS image of the kth band, $\bar{D}_k$, $\bar{B}_{Lk}$ and $\overline{(D_k)^2}$ are mean values of D_k, B_{Lk} and $(D_k)^2$, and N is the total number of pixel for each band. In addition, a quantitative way to compare the spatial quality of fused images proposed by Zhou *et al.* (1998) was adopted. It uses the correlation coefficients between the high-pass filtered fused image B^*_L and the high-pass filtered original image B_L as an index of assessing spatial quality. The ideal values of bias, standard deviation (*SD*) and root mean square error (RMSE) are 0 while smaller values represent better fusion quality in spectral feature. The ideal values of the correlation coefficient are 1. These statistical indices were calculated with each spectral band respectively.

It can be discerned that for all methods, the bias, standard deviations, and RMSE are small. The correlation coefficients are all above 0.8. With a careful examination, ISVR method generally performed the best when compared with PCS, Zhang's SVR, and GS method.

The third criterion was also adopted by visually comparing colour composites of the sets B^*_L and B_L, while it is also represented by ERGAS (Wald 2002), after its name in French '*erreur relative globale adimensionnelle de synthèse*' (dimensionless global relative error of synthesis):

$$ERGAS = 100 \times \frac{h}{l} \times \sqrt{\frac{1}{k}\sum_{i=1}^{k}\frac{(RMSE_k)^2}{M_k^2}} \tag{10}$$

in which $\frac{h}{l}$ is the ratio between pixel size of Pan and MS, e.g. 1/4 for IKONOS, M_k is the mean (average) radiance of the kth band, and K is the number of bands. An ERGAS greater than 3 corresponds to fused products of low quality, while an ERGAS lower than 3 indicates satisfactory quality (table 4), smaller ERGAS value means better quality. The colour composites of figures 2 and 3, and ERGAS in tables 4 and 5 indicate that ISVR also performs the best when compared with PCS, Zhang's SVR and GS method.

Table 5. The ERGAS index for various methods.

	ISVR	PCS	Zhang's SVR	GS
Image set 1	1.7	2.2	1.9	2.0
Image set 2	1.9	2.3	2.0	2.2

5. Conclusions

In summary, the ISVR method achieved an improved spectral and spatial quality of the fused image than PCS, Zhang's modified SVR and GS method. However, the generality of ISVR in a variety of practical applications still needs to be investigated. It has been noted that with the current ISVR method, only the visible and NIR bands are adopted for simulating the panchromatic band. That is to say, the ISVR method can only be applied to fusion of the visible and NIR bands. Further work needs to be done to investigate the suitability of the ISVR in a variety of applications.

Acknowledgements
This research was supported by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry. It was also supported by a Texasview remote sensing consortium grant awarded to Le Wang.

References

Aiazzi, B., Alparone, L., Baronti, S. and Garzelli, A., 2002, Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Transaction on Geoscience and Remote Sensing*, **40**, 2300–2312.

Chavez, P.S., Sildes, S.C. and Anderson, J.A., 1991, Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic. *Photogrammetric Engineering and Remote Sensing*, **57**, 295–303.

Munechika, C.K., Warnick, J.S., Salvaggio, C. and Schott, J.R., 1993, Resolution enhancement of multispectral image data to improve classification accuracy. *Photogrammetric Engineering and Remote Sensing*, **59**, 67–72.

Pohl, C. and van Genderen, J.L., 1998, Multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing*, **19**, 823–854.

Ranchin, T., Aiazzi, B., Alparone, L., Baronti, S. and Wald, L., 2003, Image fusion—the ARSIS concept and some successful implementation schemes. *ISPRS Journal of Photogrammetry and Remote Sensing*, **58**, 4–18.

Ranchin, T. and Wald, L., 2000, Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation. *Photogrammetric Engineering and Remote Sensing*, **66**, 49–61.

Suits, G., Malila, W. and Weller, T., 1988, Procedures for using signals from one sensor as substitutes for signals of another. *Remote Sensing of Environment*, **25**, 395–408.

Wald, L., 2002, *Data Fusion: Definitions and Architectures-Fusion of Images of Different Spatial Resolutions* (Paris, France: Les Presses, Ecole des Mines de Paris).

Wald, L., Ranchin, T. and Mangolini, M., 1997, Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images. *Photogrammetric Engineering and Remote Sensing*, **63**, 691–699.

ZHANG, Y., 1999, A new merging method and its spectral and spatial effects. *International Journal of Remote Sensing*, **20**, 2003–2014.

ZHANG, Y., 2001, Detection of urban housing development by fusing multisensor satellite data and performing spatial feature post-classification. *International Journal of Remote Sensing*, **22**, 3339–3355.

ZHANG, Y., 2004, Understanding image fusion. *Photogrammetric Engineering & Remote Sensing*, **70**, 657–661.

ZHOU, J., CIVCO, D.L. and SILANDER, J.A., 1998, A wavelet transform method to merge Landsat TM and SPOT panchromatic data. *International Journal of Remote Sensing*, **19**, 743–757.