Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

A two-step Pansharpening of ETM+ TIR image based on SFIM and neural network regression

Min Han, Wei Yao
School of Electronic and Information Engineering
Dalian University of Technology
Dalian, China
minhan@ dlut.edu.cn

***Abstract*—A two-step approach to enhance the resolution of remote sensing thermal infrared (TIR) images is proposed in this paper. For difference in imaging principles between TIR image and optical images, traditional image fusion techniques, such as component substation and MRA methods will not be proper. In our study, we use Extreme Learning Machine (ELM) to regress the relationship between TIR image and optical images, then pansharpened multi spectral images are inputted to the already trained ELM network to produce TIR image at resolution of the panchromatic image. Since the approach considers directly about radiance values in a TIR image, the result can be conveniently used in physical applications, for example, creating more precise temperature distribution of ground surface.**

***Keywords*—pansharpening, image fusion, ELM, ETM+, SFIM**

I. INTRODUCTION

The spatial resolution is one key characteristic of remote sensing data, and it directly impacts the results of the processing and the potential applications. By merging different sources of information, image fusion techniques can actually enhance the available resolutions. When the fusion is between a panchromatic (Pan) image, usually a single wide spectral-band image acquired across the visible and possibly near-infrared wavelengths,[1] and some multi spectral (MS) images with a lower spatial resolution to obtain synthetic MS images, it's called pansharpening. The Pan image is usually acquired with the maximum resolution allowed by the imaging sensor and the data link throughout, and hence contains detailed geometric features, while the MS data are acquired with a coarser spatial resolution (typically two or four times lower) as well as a better spectral resolution. So the purpose of pansharpening is to achieve high spectral and spatial resolutions at the same time.

Several spaceborne sensors, such as TM/ETM+, SPOT, Ikonos, and QuickBird, provide Pan image along with MS images, and a variety of pansharpening techniques take advantage of the complementary characteristics of spatial and spectral resolutions of the data have been founded [1]. Among those pansharpening techniques, component substitution (CS) methods [2] are attractive because they are fast and easy to implement. When exactly three MS bands are concerned, the most widely used CS fusion method is based on the intensity–hue–saturation (IHS) transformation [3]. The spectral bands are resampled and coregistered to the Pan image before the IHS transformation is applied. The smooth intensity component is substituted with the highresolution Pan and transformed back to the spectral domain via the inverse IHS transformation. Multiresolution analysis (MRA) is the alternative; it provides effective tools like wavelets and Laplacian pyramids to help carry out image fusion tasks. MRA-based fusion requires the definition of a model establishing how the missing high-pass information is to be extracted from the Pan image and then injected into the MS bands [4-5].

These methods are considered as methods of visual improvement of images by the integration of the spatial properties of high-resolution images. Sometimes, when the fusion of remote sensing images is oriented to some quantitative applications, such as regional hydrological models and carbon cycle models, these pansharpening methods are not competent, because some typical short comings of recent image fusion techniques are also exiting in these pansharpening methods.

(1) Spectral distortion: Although these current methods resemble an improvement of the resolution, they are in fact limited to improvement of visualization because the generated spectral distortions are significant. These current fusion methods largely disregard important spectral characteristics of specific satellite sensors; therefore, no consistent and color preserving results can be achieved.

(2) Poorly physic-based: More over, original radiance values represented by pixel brightness in images are seldom taken into account in those pansharpening methods, and values of digit number (DN) in those synthesized MS images can hardly transferred back into values of radiance that those MS images originally stand for.

(3) Improper for TIR: In existing pansharpening methods, the enhancement of thermal infrared (TIR) image is rarely concerned because the imaging physics of TIR image is quite different from Pan and MS images that make those methods fusing Pan and MS invalid for TIR.

Since our research focus is on the application of remote sensing and image processing techniques in a regional

This research is supported by the project (60674073) of the National Nature Science Foundation of China, the project (2006BAB14B05) of the National Key Technology R&D Program of China and the project (2006CB403405) of the National Basic Research Program of China (973 Program). All of these supports are appreciated.

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

hydrological model, we devoted to find a way to deal with these short comings listed above which hinder a more extensive application of remote sensing and image processing techniques. And the following method is just what we have achieved so far.

II. METHOD

A two step image fusion approach is proposed here to improve the spatial resolution of TIR image in ETM+ remote sensing data set. The first step is just more like a traditional pansharpening operation, and it is based on a simple spectral preserve fusion technique: the Smoothing Filter-based Intensity Modulation (SFIM) method [6]. SFIM is made use of to raise the spatial resolution of ETM+ MS images to the level of the resolution of ETM+ Pan image. Then in the second step, the synthesized MS images are used to generate high resolution TIR image according to some regression relationship between MS and TIR data. The regression relationship is obtained from original TIR image and spatially degraded MS images (resampled to scale of TIR image). Extreme Learning Machine, which is a newly developed neural network algorithm, is used as the regressor.

A. Pansharpening MS Images using SFIM

SFIM is a partially physic-based pansharpening method which can generate resolution enhanced MS images with nearly none spectral distortion. The core theory of SFIM is based on the imaging principle of reflective remote sensing images: DN values of a optical image in a reflective spectral band is mainly determined by two factors, the solar radiation impinging on the land surface $E(\lambda)$, and the spectral reflectance of the land surface $\rho(\lambda)$, thus, DN value of a certain pixel can be expressed as

$$DN(\lambda) = \rho(\lambda)E(\lambda) \tag{1}$$

Let $DN(\lambda)_{\text{low}}$ represent a DN value in a lower resolution image of spectral band λ and $DN(\gamma)_{\text{high}}$ the DN value of the corresponding pixel in a higher resolution image of spectral band γ, the SFIM method is defined as

$$\begin{aligned} DN(\lambda)_{\text{sim}} &= \frac{DN(\lambda)_{\text{low}} DN(\gamma)_{\text{high}}}{DN(\gamma)_{\text{mean}}} \\ &= \frac{\rho(\lambda)_{\text{low}} E(\lambda)_{\text{low}} \rho(\gamma)_{\text{high}} E(\gamma)_{\text{high}}}{\rho(\gamma)_{\text{low}} E(\gamma)_{\text{low}}} \\ &\approx \rho(\lambda)_{\text{low}} E(\lambda)_{\text{low}} \approx DN(\lambda)_{\text{high}} \end{aligned} \tag{2}$$

Where $DN(\lambda)_{\text{sim}}$ is the simulated higher resolution pixel corresponding to $DN(\lambda)_{\text{low}}$ and $DN(\gamma)_{\text{mean}}$ the local mean of $DN(\gamma)_{\text{high}}$ over a neighborhood equivalent to the resolution of $DN(\lambda)_{\text{low}}$. Two reasonable approximations are implied in SFIM that $E(\gamma) = E(\lambda)$ and $\rho(\lambda)_{\text{high}} = \rho(\lambda)_{\text{low}}$, thus the conclusion of $DN(\lambda)_{\text{sim}} \approx DN(\lambda)_{\text{high}}$ is achieved which means simulated image produced using SFIM can be an effective approximation for a practically acquired image in a higher resolution.

Though the factorization of DN values makes SFIM more reasonable and less arbitrary than most existing pansharpening methods, it's still not rigorous enough for physical applications. As a matter of fact, it's more proper to decompose the actual radiance value inversed from DN value as the product of solar radiation and surface reflectance, that is

$$R(\lambda) = \rho(\lambda)E(\lambda) \tag{3}$$

Where $R(\lambda)$ stands for the radiance from the land surface in the spectral band λ. Then SFIM can be rewritten as

$$R(\lambda)_{\text{sim}} = \frac{R(\lambda)_{\text{low}} R(\gamma)_{\text{high}}}{R(\gamma)_{\text{mean}}} \tag{4}$$

The radiance images, expressed as MS_1~MS_6 are obtained from MS images of ETM+ in band 1~5 & 7, and the Pan image is also transferred into a radiance format. MS_1'~MS_6' in the following equation refer to resolution enhanced MS radiance images produced using SFIM.

$$\begin{bmatrix} MS_1' \\ MS_2' \\ \vdots \\ MS_6' \end{bmatrix} = r \begin{bmatrix} MS_1 \\ MS_2 \\ \vdots \\ MS_6 \end{bmatrix} = \frac{\text{Pan}}{\text{Pan}'} \begin{bmatrix} MS_1 \\ MS_2 \\ \vdots \\ MS_6 \end{bmatrix} \tag{5}$$

The ultimate purpose of our research is to raise resolution of the TIR image to the level of the Pan image; however, we can not apply SFIM to the TIR image directly because the TIR image is not in a reflective spectral band and thermal infrared radiation which is decided by emissivity and temperature of the ground surface rather than solar radiation and surface reflectance, can not be decomposed following the way of equation (3). Therefore an alternative approach is required.

Though the relevance between reflective information and thermal radiation information is indirect and subtle, the reflection and radiation behaviors are both decided by the ground surface character corresponding to each individual pixel when the solar radiation is more or less the same for a considerably large area under slight cloud cover conditions, which means there are some inner links between ground surface radiation and reflectance. So to generate thermal radiance information from reflective radiance information could be an effective way to synthesize thermal image with higher resolution.

B. Regreesion for TIR Image

Kustas et al in 2003 [7] proposed an approach to estimate subpixel surface temperature in a remote sensing image using the vegetation index-radiometric temperature relationship. The relationship between radiometric temperature (T_r) and normalized difference vegetation index (NDVI) is obtained from a linear least square regression of low resolution T_r and spatially degraded NDVI data, then the relationship and original high resolution NDVI data are utilized in a disaggregation procedure for estimating subpixel variation in surface temperature. Surface radiometric temperature T_r could be considered as an equivalent of surface thermal infrared radiance.

Though the idea of regression is inspiring, this approach is not well designed. Two main weak points exist. First, as in an ETM+ image, NDVI only concerns about information in 2 out of the 6 reflective bands, and even the only considered red band and the near infrared band values are merged into one variable, which means a big part of information in those reflective bands have been neglected. Furthermore, the least square fit is an inefficient algorithm comparing to some newly established machine learning methods such as neural network regressions, and limitations in the shapes that linear models can assume over long ranges is also a big disadvantages of linear least squares.

In our study, the relationship between all the 6 MS bands of radiance data and the thermal radiance data is regressed, while a more effective regression method, Extreme Learning Machine (ELM), is introduced to do the job [8].

ELM is a simple and efficient learning algorithm for single-hidden layer feed forward neural networks. The learning speed of ELM is extremely fast and better generalization performances than the gradient-based learning such as back propagation (BP) in most cases have also been confirmed.

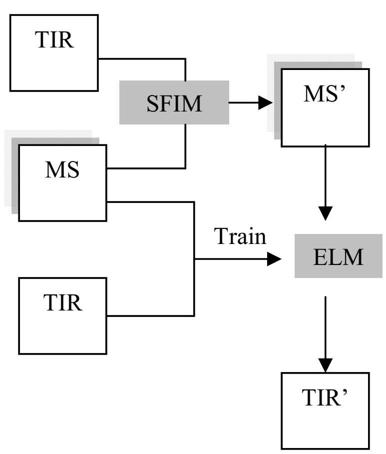

Figure 1. Flowchart of the two step pansharpening method for an ETM+ TIR image

In order to obtain a nonlinear regression relationship between radiance of MS bands and TIR band in an ETM+ image, all the 6 bands of MS images are resampled to the scale of TIR image preliminarily. Radiance values in these spatially degraded MS images and the TIR image are then used to train the ELM. When the training is done, the regression relationship between MS radiance and thermal radiance is established in the ELM network. Then, the previously obtained 15m resolution SFIM-sharpened MS images are used as inputs to the properly trained ELM network, and the TIR image produced via the network will take a resolution as high as the Pan image. Figure 1 is the flowchart of the two step pansharpening method for an ETM+ TIR image. MS' refers to pansharpened MS images using SFIM and TIR' stands for the 15m resolution TIR image synthesized via the two step approach.

C. Discusion

It has been proved by an IHS translation of the pansharpened images that SFIM do little spectral distortion [9]. Furthermore, through the whole pansharpening procedure, remote sensing images are in the form of radiance values, and results of obtained using the approach are also radiance values that can be directly used in further practical applications. As a matter of fact, radiance data in this synthesized resolution enhanced TIR image will be used in another part of our research works aiming to calculate regional evapotranspiration amount under the framework of a hydrological remote sensing model. Since this follow-up part of work is not that closely related to the topical of this script, the results are not included.

III. Experiment

The experiment is limited to ETM+ data and the whole procedure can be divided into 3 parts: the training process for the ELM network; the solo utilization of trained ELM to generated TIR image with resolution of 30m; the combined process of SFIM and ELM to do the pansharpening of the TIR image.

A. Experiment Data Sets

The ETM+ image used in our study is acquired on July 4th and September 22nd, 2002. Path and row numbers of the image are 112 and 27 respectively, corresponding to the area of Zhalong wetland in north China. A sub region with 190×220 pixels (as in MS bands) of the September 22nd image is selected, and the spatially degraded 95×110 groups of MS radiance values together with the original 95×110 TIR radiance values form the training data set for the ELM network.

To test the generalization performance of the regression relationship established in the trained ELM network, another sub region of 250×270 corresponding to a different location in the September 22nd image is selected, and radiance values for those pixels are organized as testing data set 1. An extra testing data set is also obtained from a 190×220 sub image in the July 4th image which covers the same region as the training data set, and this data set is labeled as testing data set 2.

B. Programming and Configuration

In our study, Erdas Imagine 9.0 is used to do some pretreatment on these remote sensing images and then to extract the 3 sub images containing radiance values from original MS, TIR and Pan images of DN values in ETM+ data files. And all the following programming and experiments are implemented in Matlab 2007a and run on a Pentium Dual 2-GHz machine with 2-GB random access memory.

The codes of ELM are downloaded from the homepage of the researcher who originally proposed the ELM network and more details about ELM and related researches are available at http://www3.ntu.edu.sg/home/egbhuang/.

IV. RESULTS

MS images in an ETM+ data set are at the resolution of 30m, while the Pan image and TIR image are at 15m and 60m respectively. All the image data are previously prepared in the form of radiance value.

A. Training and Testing

First, the training data set is used to train the ELM, while number of hidden layer neurons is set to be 100 and Sigmoidal function is chosen to be the activation function of the ELM network. The Root Mean Square Error (RMSE) for the training results is 0.0925. And when test the trained ELM with the testing data set 1 and testing data set 2, we get RMSEs of 0.1458 and 0.9072 respectively.

Figure 2~4 show the results of the training and testing procedures. Where the horizontal axis is the output of ELM, and the vertical axis is the true value of TIR radiance.

The training result and testing result of testing data set 1 are fairly satisfactory as shown in figure 2 and figure 3, a clear consistency between real data and model output can be observed. It can be concluded from the testing result of testing data set 1 that the regression relationship between MS radiance data and TIR radiance data obtained in one region hold true for another region in the same remote sensing image.

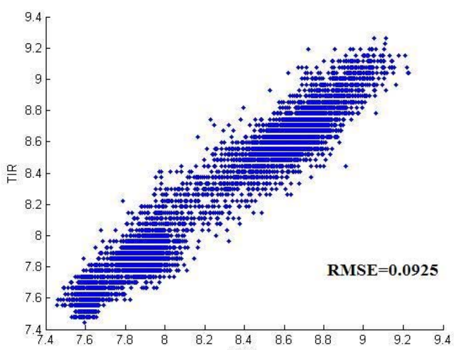

Figure 2. Training result of the ELM network

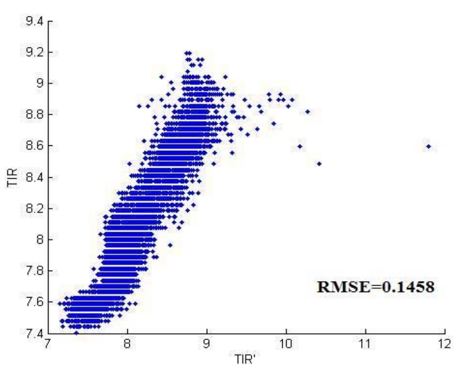

Figure 3. Testing result of the ELM network using testing data set 1

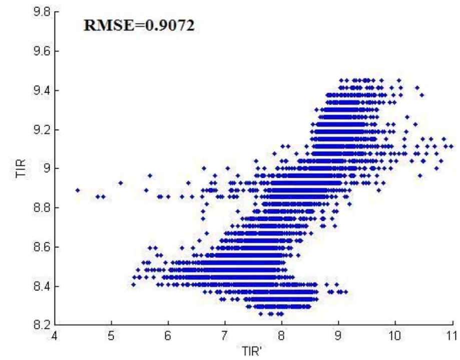

Figure 4. Testing result of the ELM network using testing data set 2

As shown in figure 4, a larger RMSE for the testing data set 2 is obtained and a worse consistence between real TIR data and ELM output is observed. Though in our assumption, the relationship between MS radiance and TIR radiance data should be universal since it only concerns about the surface condition. This difference between results and assumption could be ascribed to the atmospheric correction procedure which is not accurate enough. For a practical application, this result implies that for remote sensing images acquired on different days, the regression relationships are not the same.

B. Results of ELM Approach

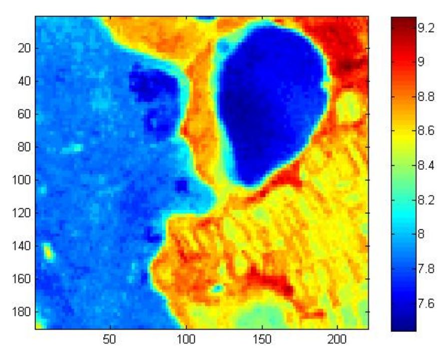

Figure 5. Original ETM+ TIR image of a subregion in a wetland area

A synthesized TIR image at the resolution of 30m is obtained when directly use the trained ELM with original MS radiance data. To assess the fidelity of the results, The synthesized TIR image is resampled back to the resolution of 60m to compare with the original TIR image. A RMSE of 0.1360 exists for the synthesized and original data set.

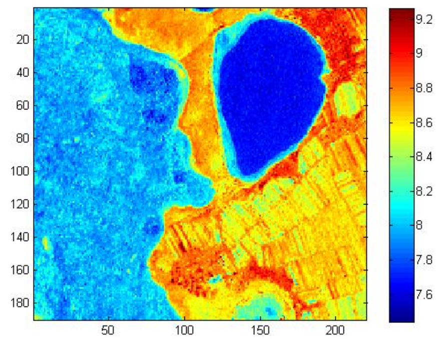

Figure 6. Pansharpened ETM+ TIR image synthersized only by the ELM network

C. Results of SFIM+ELM Approach

Using SFIM, we get resolution enhanced MS radiance images at the resolution of 15m. With these data, a 15m TIR image can be obtained from the trained ELM. The fidelity of this high resolution TIR image is also evaluated with the method mentioned above, and a larger RMSE of 0.2319 is obtained.

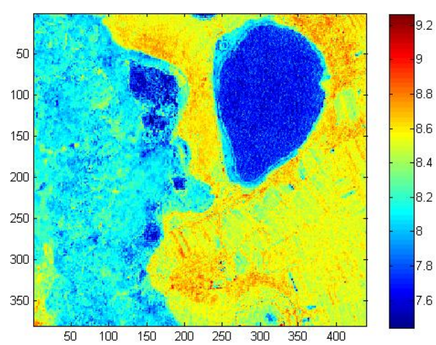

Figure 7. Futther sharpened ETM+ TIR image via the two-step approach

V. CONCLUSION

Thermal infrared images are usually coarser than MS images which is a limitation for many applications, while thermal infrared radiation is closely related with ground surface temperature and information offered by TIR images is essential for the monitoring and modeling of ground surface processions. Thus, a promotion of the TIR image quality is not only meaningful but also necessary.

A two step approach is proposed in this paper to deal with resolution enhancement of TIR images, which is usually ignored in researches of Pansharpening and image fusion techniques. The approach is well physic-based that through all the processing procedures images are in the form of radiance values.

Acceptable results have been achieved using the two-step sharpening, though the generalization performances of the method for data sets acquired in different days are not good enough. This generalization problem could be induced by different atmosphere conditions which will be taken into account in the further study of this TIR pansharpening technique.

ACKNOWLEDGMENT

The authors would like to thank professor Guang-Bin Huang for sharing the codes of ELM. We also thank China Remote Sensing Satellite Ground Station for providing ETM+ data.

REFERENCES

[1] W. Zhijun, D. Ziou, C. Armenakis, D. Li, and L. Qingquan, "A comparative analysis of image fusion methods," Geoscience and Remote Sensing, IEEE Transactions on, vol. 43, pp. 1391-1402, 2005.

[2] B. Aiazzi, S. Baronti, and M. Selva, "Improving Component Substitution Pansharpening Through Multivariate Regression of MS +Pan Data," Geoscience and Remote Sensing, IEEE Transactions on, vol. 45, pp. 3230-3239, 2007.

[3] T. Te-Ming, P. S. Huang, H. Chung-Ling, and C. Chien-Ping, "A fast intensity-hue-saturation fusion technique with spectral adjustment for IKONOS imagery," Geoscience and Remote Sensing Letters, IEEE, vol. 1, pp. 309-312, 2004.

[4] T. Ranchin and L. Wald, "Fusion of High Spatial and Spectral Resolution Images: the ARSIS Concept and Its Implementation," Photogrammetric Engineering & Remote Sensing, vol. 66, pp. 49-61, 2000 2000.

[5] M. Gonzalez-Audicana, J. L. Saleta, R. G. Catalan, and R. Garcia, "Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition," Geoscience and Remote Sensing, IEEE Transactions on, vol. 42, pp. 1291-1299, 2004.

[6] J. G. Liu, "Smoothing Filter-based Intensity Modulation: a spectral preserve image fusion technique for improving spatial details," International Journal of Remote Sensing, vol. 21, pp. 3461-3472, 2000.

[7] W. P. Kustas, J. M. Norman, M. C. Anderson, and A. N. French, "Estimating subpixel surface temperatures and energy fluxes from the vegetation index-radiometric temperature relationship," Remote Sensing of Environment, vol. 85, pp. 429-440, 2003.

[8] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," Neurocomputing, vol. 70, pp. 489-501, 2006.

[9] T.-M. Tu, Y.-C. Lee, P. S. Huang, and C.-P. Chang, "Modified smoothing-filter-based technique for IKONOS-QuickBird image fusion," Optical Engineering, vol. 45, pp. 066201-10, 2006.