FUSION OF OPTICAL AND SAR IMAGES FOR THE ENHANCEMENT OF URBAN FEATURES

D.Amarsaikhan[1], Ch.Bolorchuluun[2], N.Ganchuluun[1], M.Ganzorig[1], B.Nergui[1] and D.Enkhjargal[1]

[1]Institute of Informatics and RS, Mongolian Academy of Sciences
av.Enkhtaivan-54B, Ulaanbaatar-51, Mongolia
Tel: 976-11-453660, Fax: 976-11-458090

[2]Department of Geography, National University of Mongolia
Ikh Surguuliin gudamj-6, Ulaanbaatar-46, Mongolia
Tel/Fax: 976-11-322822

KEY WORDS: Fusion techniques, Optical, Synthetic aperture radar (SAR), Urban area

ABSTRACT: The aim of this study is to explore the performances of diverse data fusion methods for the enhancement of urban land cover features when multisource images are under consideration. For the data fusion, some ordinary and advanced data fusion techniques are used and the results are compared. As a test site, Ulaanbaatar, the capital city of Mongolia is selected and as data sources multichannel Landsat ETM(+) image and Envisat SAR data are used.

1. INTRODUCTION

Over the years, the image fusion has become a very valuable approach for the integration of multi-sensor data sets. Very often it is used to produce an image with an improved spatial resolution. The most common situation is represented by a pair of images where the first acquired by a multispectral sensor has a pixel size greater than the pixel size of the second image acquired by a panchromatic sensor. Combining these images, fusion produces a new multispectral image with a spatial resolution equal to the panchromatic one. It has been found that the images acquired at different ranges of electro-magnetic spectrum provide unique information when they are integrated. In case of RS data sets, three different fusions such as fusion of optical data with optical data, fusion of microwave data with microwave data and fusion of optical and microwave data sets can be conducted (Amarsaikhan *et al.* 2012).

Among three different fusions, the image fusion based on the integration of optical and microwave data sets, is being efficiently used for the interpretation, enhancement and analysis of different land surface features. It is evident that a combined use of the optical and SAR images will have a number of advantages because a specific feature which is not seen on the passive sensor image might be seen on the microwave image and vice versa because of the complementary information provided by the two sources (Amarsaikhan *et al.* 2007). Now image fusion based on the integration of multispectral optical and multifrequency microwave data sets is being efficiently used for different analyses. Many authors have proposed and applied different techniques to combine optical and SAR images and they all judged that the results were better (Van Genderen 1998, Amarsaikhan and Douglas 2004, Ehlers *et al.* 2008, Zhang 2010).

The aim of this study is to evaluate and compare such ordinary and advanced data fusion techniques as Brovey transform, Gram-Schmidt method, principal component analysis (PCA), modified intensity-hue-saturation method (IHS), wavelet-based method and Elhers fusion for the enhancement of spectral variations of different urban land cover types which represent a unique overview of modern and nomadic cultures. As a test site, Ulaanbaatar, the capital city of Mongolia has been selected. As data sources, multispectral Landsat ETM(+) data as well as Envisat radar image of 2011 have been used. The actual analysis was carried out using PC-based ERDAS Imagine 10.1 and ENVI 4.9.

2. TEST SITE AND DATA SOURCES

As a test site, Ulaanbaatar, the capital city of Mongolia has been selected. Ulaanbaatar is situated in the central part of Mongolia, on the Tuul River, at an average height of 1350m above sea level and currently has about 1.3 million inhabitants (NSO, 2013). The city is surrounded by the mountains which are spurs of the Khentii Mountain Ranges. Founded in 1639 as a small town named Urga, today it has prospered as the main political, economic, business, scientific and cultural center of the country.

The study area chosen for the current study covers mainly the central, eastern and southern parts and is characterized by such classes as built-up area, ger (Mongolian national dwelling) area, green area, soil and water. Figure 1 shows a Landsat image of the test site, and one can see what land cover types are available.

Figure 1. Landsat image of the selected part of Ulaanbaatar city.

In the present study, as data sources, a Landsat ETM(+) image of 31 July 2010 and Envisat data of January 2011 have been used. The Landsat ETM(+) data has seven multispectral bands (B1: 0.45–0.52µm, B2: 0.52–0.60µm, B3: 0.63–0.69µm, B4: 0.76–0.90µm, B5: 1.55–1.75µm, B6: 10.40–12.50µm and B7: 2.08–2.35µm). The spatial resolution is 30 m for the reflective bands, while it is 60 m for the thermal band. In this study, channels 2,3,4,5,7 have been used. The Envisat is a European satellite carrying all weather free polarimetric radar. In our study, the Envisat HH polarisation image with a spatial resolution of 25m has been used. In addition, a topographic map of 1969 and 1984, scale 1:50 000, and a general urban planning map were available.

3. COREGISTRATION OF MULTISENSOR IMAGES AND SPECKLE SUPPRESSION OF SAR IMAGE

At first, a Landsat image was georeferenced to a UTM map projection using 12 ground control points (GCPs) defined from a topographic map of the study area, scale 1:100,000. The GCPs have been selected on well delineated crossings of roads and city building corners. For the transformation, a second-order transformation and nearest-neighbour resampling approach were applied and the related root mean square (RMS) error was 0.85 pixel. Then, the SAR image was georeferenced and its coordinates were transformed to the coordinates of the rectified optical image. In order to correct the SAR image, 15 more regularly distributed GCPs were selected from different parts of the microwave image. For the actual transformation, a second-order transformation was used. As a resampling technique, the nearest-neighbour resampling approach was applied and the related RMS error was 1.16 pixel.

The microwave images have a granular appearance due to the speckle formed as a result of the coherent radiation used for radar systems. Therefore, the speckle reduction is a very important step in further analysis. In addition, the analysis of the radar images must be based on the techniques that remove the speckle effects while considering the intrinsic texture of the image frame (Serkan *et al.* 2008). Considering these requirements, a 3x3 gammamap filter has been chosen. As could be seen, in the output SAR image, the speckle noise was reduced with very low degradation of the textural information.

4. FUSION ANALYSIS OF MULTISOURCE IMAGES

Generally, image fusion can be performed at pixel, feature and decision levels (Pohl and Van Genderen 1998). In this study, data fusion has been performed at a pixel level and some common and more complex techniques were compared. Each of these techniques is briefly discussed below.

Brovey transform: This is a simple numerical method used to merge different digital data sets. The algorithm based on a Brovey transform uses a formula that normalises multispectral bands used for a red, green, blue colour display

and multiplies the result by high resolution data to add the intensity or brightness component of the image (Vrabel 1996). The formulae used for the Brovey transform can be described as follows:

$$\text{Red} = \frac{\text{Band1}}{\sum_{i=1}^{n} \text{Bandn}} * \text{High Resolution Band}$$

$$\text{Green} = \frac{\text{Band2}}{\sum_{i=1}^{n} \text{Bandn}} * \text{High Resolution Band}$$

$$\text{Blue} = \frac{\text{Band3}}{\sum_{i=1}^{n} \text{Bandn}} * \text{High Resolution Band}$$

PCA: The most common understanding of the PCA is that it is a data compression technique used to reduce the dimensionality of the multidimensional datasets or bands. The bands of the PCA data are noncorrelated and are often more interpretable than the source data. The process is easily explained if we consider a two dimensional histogram which forms an ellipse. When the PCA is performed, the axes of the spectral space are rotated, changing the coordinates of each pixel in spectral space. The new axes are parallel to the axes of the ellipse. The length and direction of the widest transect of the ellipse are calculated using a matrix algebra. The transect, which corresponds to the major axis of the ellipse, is called the first principal component of the data. The direction of the first principal component is the first eigenvector, and its length is the first eigenvalue. A new axis of the spectral space is defined by this first principal component. The second principal component is the widest transect of the ellipse that is perpendicular to the first principal component. As such, the second principal component describes the largest amount of variance in the data that is not already described by the first principal component. In a two-dimensional case, the second principal component corresponds to the minor axis of the ellipse. In *n* dimensions, there are *n* principal components (Richards and Jia, 1999).

Gram-Schmidt fusion method: Gram-Schmidt process is a procedure which takes a non-orthogonal set of linearly independent functions and constructs an orthogonal basis over an arbitrary interval with respect to an arbitrary weighting function. In other words, this method creates from the correlated components non- or less correlated components by applying orthogonalization process (Karathanassi *et al.* 2008).

In any inner product space, we can choose the basis to work. It often simplifies the calculations to work in an orthogonal basis. Let us suppose that K = $\{v_1, v_2, \ldots, v_n\}$ is an orthogonal basis for an inner product space V. Then it is a simple matter to express any vector $w \in V$ as a linear combination of the vectors in K:

$$\omega = \frac{\langle w, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle w, v_2 \rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle w, v_n \rangle}{\|v_n\|^2} v_n$$

Given an arbitrary basis $\{u_1, u_2, \ldots, u_n\}$ for n-dimensional inner product space V, the Gram-Schmidt algorithm constructs an orthogonal $\{v_1, v_2, \ldots, v_n\}$ for V and the process can be described as follows:

$v_1 = u_1$

$v_2 = u_2\text{-project}_{w1}\, u_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1$

$v_3 = u_3\text{-project}_{w2}\, u_3 = u_3 - \frac{\langle u_3, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_3, v_2 \rangle}{\|v_2\|^2} v_2$

$v_4 = u_4\text{-project}_{w3}\, u_4 = u_4 - \frac{\langle u_4, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_4, v_2 \rangle}{\|v_2\|^2} v_2 - \frac{\langle u_4, v_3 \rangle}{\|v_3\|^2} v_3.$

Where:

w1-space spanned by v_1
$\text{project}_{w1}\, u_2$ is the orthogonal projection of u_2 on v_1
w2-space spanned by v_1 and v_2
w3-space spanned by v_1, v_2 and v_3.

This process continues up to v_n. The resulting orthogonal set $\{v_1, v_2, \ldots, v_n\}$ consists of n-linearly independent vectors in V and forms an orthogonal basis for V.

The modified IHS: This method has a vast improvement over traditional IHS for fusing satellite imagery that differs noticeably in spectral response. It allows combining single band panchromatic data with multispectral data, resulting in an output with both excellent detail and a realistic representation of original multispectral scene colors. The modified IHS method is designed to produce an output that approximates the spectral characteristics of the input multispectral bands while preserving the spatial integrity of the panchromatic data. The technique works by assessing the spectral overlap between each multispectral band and the high resolution panchromatic band and weighting the merge based on these relative wavelengths (ERDAS 1999).

Wavelet-based fusion: The wavelet transform decomposes the signal based on elementary functions, that is the wavelets. By using this, an image is decomposed into a set of multi-resolution images with wavelet coefficients. For each level, the coefficients contain spatial differences between two successive resolution levels. The wavelet transform can be expressed as follows:

$$WT(f)(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t)\varphi\left(\frac{t-b}{a}\right) dt$$

Where:

a-scale parameter
b-translation parameter.

Practical implementation of the wavelet transform requires discretisation of its translation and scale parameters. In general, a wavelet-based image fusion can be performed by either replacing some wavelet coefficients of the low-resolution image by the corresponding coefficients of the high-resolution image or by adding high resolution coefficients to the low-resolution data (Pajares and Cruz, 2004). In the present study, 'Wavelet Resolution Merge' tool of Erdas Imagine was used and the algorithm behind this tool uses biorthogonal transforms.

Elhers fusion: This is a fusion technique used for the spectral characteristics preservation of multitemporal and multi-sensor data sets. The fusion is based on an IHS transformation combined with filtering in the Fourier domain and the IHS transform is used for optimal colour separation. As the spectral characteristics of the multispectral bands are preserved during the fusion process, there is no dependency on the selection or order of bands for the IHS transform. Unlike the standard approach, the method is extended to include more than 3 bands using multiple IHS transforms until the number of bands is fulfilled. A subsequent Fourier transform of the intensity component and the panchromatic image allows an adaptive filter design in the frequency domain. By the use of the fast Fourier transform (FFT) techniques, the spatial components to be enhanced or suppressed can be directly accessed. The intensity spectrum is filtered with a low pass filter (LP) whereas the panchromatic spectrum is filtered with an inverse high pass filter (HP). After filtering, the images are transformed back into the spatial domain with an inverse FFT and added together to form a fused intensity component with the low-frequency information from the low resolution multispectral image and the high-frequency information from the panchromatic image. This new intensity component and the original hue and saturation components of the multispectral image form a new IHS image. As the last step, an inverse IHS transformation produces a fused RGB image (Ehlers *et al.* 2008).

At the beginning, the above mentioned fusion approaches have been applied to the combined optical and SAR images. To obtain good colour images that can illustrate spectral and spatial variations of the classes of objects on the images, the fused images have been visually compared. On the Brovey transformed image, the built-up and ger areas had similar appearances and it somehow reflected the characteristics of the SAR image. Also, on the image water and soil classes as well as green vegetation and forest had comparable appearances. Therefore, it is possible to judge that the result of the Brovey transform is not suitable for the spectral separation of the available land cover classes. The Gram-Schmidt method created a good color image, because it gave some separations among the green area, soil and water classes. However, on the image the built-up and ger areas had a bit similar appearances, tough they were texturally distinguishable.

In the case of the PCA, the fused image showed a better result compared to some other combinations. The PC image created by combination of the first three PCs (they contained more than 98% of the overall variance) demonstrated very much the characteristics of both optical and radar images, because, in the PC1 the Envisat image had a very high negative loading, while in the PC2 and in PC3, infrared bands of the Landsat had high loadings. The modified IHS and wavelet-based methods gave superior results and the images looked better than any other images obtained by other fusion methods. On these images, all available five classes could be distinguished by their spectral properties. Moreover, it could be seen that some textural information has been added for differentiation between the classes: built-up area and ger area. In the case of the Elhers fusion, the image demonstrated the worst result compared to all other combinations. This image had a blurred appearance due to speckle noise and it was

very difficult to observe the separation among the available classes. Figure 2 shows the comparison of the images obtained by the applied fusion methods.

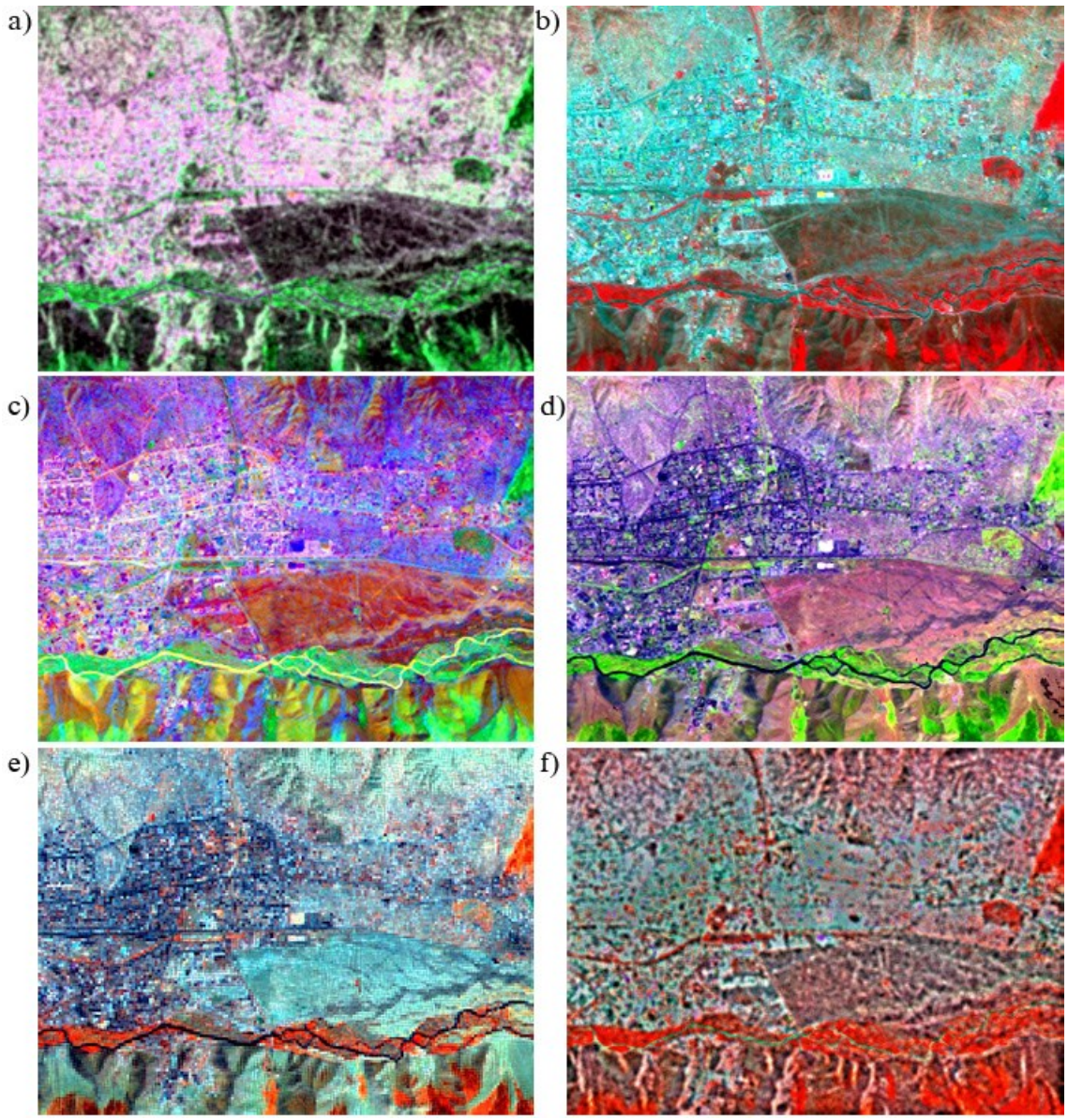

Figure 2. Comparison of the fused images of Landsat and Envisat SAR:
(a) Brovey transformed image; (b) the image obtained by Gram-Schmidt fusion;
(c) PC image; (d) the image obtained by the modified IHS method;
(e) the image obtained by wavelet-based fusion; (f) the image obtained by Elhers fusion.

5. CONCLUSIONS

The overall idea of the research was to compare the performances of different data fusion techniques for the enhancement of different urban features related to Ulaanbaatar city, Mongolia. For the data fusion, Brovey transform, Gram-Schmidt method, PCA, modified IHS, wavelet-based method and Elhers fusion were used. Although, fusion techniques showed different results, detailed analysis of each image revealed that the images obtained by the modified IHS and wavelet-based method gave superior images in terms of the spectral and spatial separation among different urban features.

6. REFERENCES

Amarsaikhan, D. and Douglas, T., 2004. Data fusion and multisource data classification. International Journal of Remote Sensing, 17, pp.3529-3539.

Amarsaikhan, D., Ganzorig, M., Ache, P. and Blotevogel, H., 2007. The Integrated Use of Optical and InSAR Data for Urban Land Cover Mapping, International Journal of Remote Sensing, 28, pp.1161-1171.

Amarsaikhan, D., Ganzorig, M., Saandar, M., Blotevogel, H.H., Egshiglen, E., Gantuya, R., Nergui, B. and Enkhjargal, D., 2012. Comparison of multisource image fusion methods and land cover classification, International Journal of Remote Sensing, Vol.33, pp.2532-2550.

Ehlers, M., Klonus, S. and Åstrand, P.J., 2008. Quality Assessment for Multi-sensor Multi-date Image Fusion, CD-ROM Proceedings of ISPRS Congresses, Beijing, China.

ERDAS 1999. Field Guide, 5th edn (Atlanta, Georgia: ERDAS, Inc.).

Karathanassi, V., Kolokousis, P. and Ioannidou, S., 2007. A comparison study on fusion methods using evaluation indicators. International Journal of Remote Sensing, 28, pp.2309 – 2341.

National Statistical Office (NSO) of Mongolia, 2013. Mongolian Statistical Yearbook 2010, Ulaanbaatar, Mongolia.

Pajares, G and Cruz, J.M., 2004. A wavelet-based image fusion, Pattern Recognition, Volume 37, Issue 9, September 2004, pp.1855-1872.

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

Richards, J. A. and Jia, S., 1999. Remote Sensing Digital Image Analysis—An Introduction, 3rd edn (Berlin: Springer-Verlag).

Serkan, M., N. Musaoglu, H. Kirkici and C. Ormeci, 2008. Edge and fine detail preservation in SAR images through speckle reduction with an adaptive mean filter, International Journal of Remote Sensing, Vol. 29, No. 23, pp.6727 – 6738.

Vrabel, J., 1996. Multispectral imagery band sharpening study. Photogrammetric Engineering and Remote Sensing, 62, pp.1075-1083.

Zhang, J., 2010. Multi-source remote sensing data fusion: status and trends. International Journal of Image and Data Fusion, 1, pp.5 – 24.