QUALITY ASSESSMENT OF FUSION METHODS FOR HIGH RESOLUTION IMAGES

Meenakshisundaram, V. [a,*] & Couloigner, I. [a]

[a] Department of Geomatics Engineering, University of Calgary, Alberta-T2N1N4 – (vmeenaks, icouloig)@ucalgary.ca

KEY WORDS: Quickbird, Data fusion, Quantitative quality assessment

ABSTRACT:

High resolution panchromatic (PAN) images provide better spatial quality compared to multispectral (MS) images. However MS images provide better spectral quality compared to PAN images. Object recognition and extraction of urban features require methods to integrate complementary data sets such as PAN, MS or Lidar data. Many pixel-based image fusion techniques have been developed to combine the spatial and spectral characteristics of PAN and MS images in synthesized MS images However it is important to evaluate the spatial and spectral quality of these synthesized MS images. In this communication, we will present the results of three pixel-based fusion methods for Quickbird PAN and MS images. The degree to which the spectral content of the original image is preserved depends on the fusion method used. The standard method IHS (Intensity-Hue-Saturation) and wavelet transform methods have been evaluated based on statistical criteria. Although there are different criteria to evaluate the synthesized images, the assessment of their multispectral quality remains a challenge. Different features tend to have different spectral distortions in different bands. In this communication, apart from the usual statistics, we have selected different clusters representing different classes to evaluate the spectral distortion. The error in the synthesis (Root Mean Square Error) has been calculated for each cluster for evaluating the quality of the fusion.

1. INTRODUCTION

Availability of high spatial resolution satellite images, advancements in digital image processing and computer technology motivate research in automation of urban feature extraction. Due to the high complexity of urban environment, automatic feature extraction is not yet operational. The first step in feature extraction is to extract edges and regions from the images using different segmentation methods. Edges represent the object boundaries and PAN is more suitable for this purpose. Segmentation of MS images result in homogeneous regions that can be used to identify objects. In remote sensing, MS images are often used for mapping as they can be automatically classified with minimal human intervention. In the field of pattern recognition, PAN images are preferred as they have higher spatial resolution compared to the MS images one. For urban mapping, we need complementary data sets (e.g. PAN, MS, or Lidar) like PAN for feature extraction, MS for object recognition (Couloigner *et al.*, 1998) and Lidar data for surface description (Csathó *et al.*, 2003). Since PAN and MS have different spatial resolutions, the fusion at feature level may be difficult. Instead, we can use pixel-based fusion methods to combine the characteristics of PAN and MS images to obtain MS images with good spatial as well as good spectral information. This may considerably reduce the processing required at feature level when feature primitives (edges and/or regions) are extracted from the same synthesized MS images. While this integration may be useful, it is necessary to ensure that the resultant synthesized images have the same spectral content than the original MS images. Indeed classification procedures depend on spectral signatures. There are several methods for fusing PAN and MS images. In this paper, we consider three data fusion methods, namely IHS (Intensity-Hue-Saturation), IHS with wavelet (IHSW) and ARSIS ("Amélioration de la Résolution Spatiale par Injection of Structures") methods, for analyzing the spectral quality of the synthesized Quickbird images.

2. DATA AND METHODOLOGY

The Quickbird PAN image has a spatial resolution of 0.7 m obtained in a spectral range of 0.45 to 0.90 µm and the MS images have a spatial resolution of 2.8 m obtained in four spectral ranges of (blue) 0.45-0.52 µm, (green) 0.52-0.60 µm, (red) 0.63-0.0.69 µm, and (NIR) 0.76-0.90 µm. Our case study consists of orthorectified Quickbird data sets over the City of Fredericton, New Brunswick, Canada. The data sets are images over urban area and consist of different features like road, building, parking, tree, grass, and so on (see Figure 1). The PAN image has a dimension of 1024 x 1024 pixels and the MS subsets have a dimension of 256 x 256 pixels. A subset of the green band is shown in Figure 1. This subset is used to evaluate the performance of the IHS, IHSW and ARSIS methods.

Figure 1. Orthorectified Quickbird subset in the green band

The IHS method is based on the human color perception parameters. It separates the spatial (I) and spectral (H, S) components of a RGB image. Intensity (I) refers to the total brightness of the color. Hue (H) refers to the dominant wavelength and Saturation (S) refers to the purity of the color relative to gray. This method involves the replacement of the intensity (I) component by a high spatial resolution PAN image and then reversing to the RGB color space. The PAN image is contrast stretched or histogram matched to the I component it replaces. The model for tranforming the RGB space to the IHS space that has been used in this study is given in Pohl and Genderen (1998).

For the IHSW method, the modifications of the IHS method proposed by Nunez *et al*. (1999) was used. Instead of replacing the I component by a PAN image, the high frequency components (i.e. the details) from PAN are modeled using the à-trous algorithm (Dutilleux, 1987) with a B_3-spline scaling function and injected into the I component. Finally, the components are transformed back to the RGB space.

For the ARSIS method, a model is built between the wavelet coefficients of PAN and of MS images at the same scale. This model is then used to synthesize the MS image at the same spatial resolution than the PAN one. A complete description of the ARSIS method can be found in Ranchin *et al*. (2003).
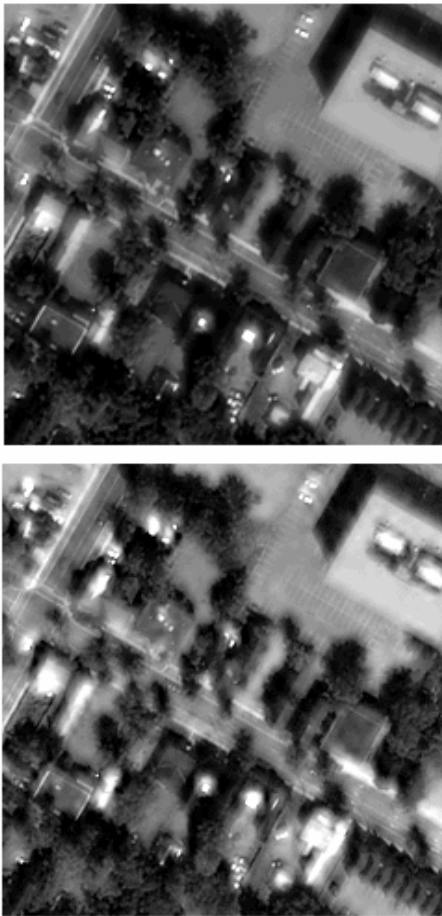

Figure 2. Synthesized images by
Top: ARSIS method; Bottom: IHSW method

3. RESULTS AND ANALYSIS

A subset of the synthesized green band resulting from the ARSIS and the IHSW methods is shown in Figure 2. This figure shows that the ARSIS method provides a better visual quality compared to the IHSW method. However visual assessment of the synthesized images does not provide information about the spectral preservation of the images. Therefore, a statistical analysis has been performed.

3.1 Statistical Analysis

Different combinations of the input images were selected for the IHS method. The numbers 4, 3, 2 and 1 refer to the NIR, Red, Green and Blue bands respectively. We are only presenting the statistics for the 431 and 421 combinations for the IHS and IHSW methods in Table 3. There are many different models in the ARSIS method and only the results of the MSM3M2 (Multi-Scale Model with 3 x 3 filter and Model 2 (M2) IBSM and HRIBSM) is given.

The original MS images have been used as reference to compute the statistics. For analysis purpose, the synthesized MS images (MS_{Syn}) have been resampled to the original resolution (2.8 m) using a bicubic interpolation. They are denoted by MS_{Syn}^{*}.

	IHS		IHSW		ARSIS
	431	**421**	**431**	**421**	**MSM3M2**
Bias					
Blue	-0.27	-0.28	0	0	0.26
Green	-	-0.21	-	0	0.21
Red	-0.33	-	0	-	0.33
NIR	-0.16	-0.17	0	0	0.07
Std dev					
Blue	23.8	25.46	9.73	10.41	14.78
Green	-	19.25	-	7.86	19.53
Red	28.93	-	11.82	-	26.25
NIR	14.33	15.33	5.84	6.25	23.19
Var					
Blue	-4.09	-3.92	-14.75	-16.43	10.3
Green	-	11.06	-	-7.27	10.03
Red	7.96	-	-8.06	-	9.8
NIR	4.62	2.89	-3.83	-4.16	9.76
Corr coef					
Blue	0.76	0.73	0.96	0.96	0.9
Green	-	0.91	-	0.99	0.9
Red	0.89	-	0.98	-	0.91
NIR	0.96	0.96	0.99	0.99	0.9
ERGAS	5.79	5.11	2.36	2.09	5.34

Table 3. Statistics: Bias, Standard deviation, Variance, Correlation Coefficient in percentage and ERGAS value for the different synthesized MS bands.

The statistics computed to compare the spectral quality of the synthesized MS images resulting from the three applied methods are presented in Table 3. Bias is the difference between the mean of MS and MS_{Syn}^{*}. It is given in percentage relative to the mean of MS. The standard deviation (Std dev) is

the standard deviation of the difference image ($MS - MS^*_{Syn}$) given in percentage relative to the mean of MS. The variance (Var) is the difference between the variance of MS and MS^*_{Syn} given in percentage relative to the variance of MS. Correlation coefficient (Corr coef) is the correlation coefficient between MS and MS^*_{Syn}. The ideal value is zero for bias, standard deviation and variance, and one for correlation coefficient. A positive variance indicates a loss of information. A negative variance indicates added information, *i.e.* detail information not present in the original MS images is present in the synthesized MS^*_{Syn} images. ERGAS (from the French "Erreur Relative Globale Adimensionnelle de Synthèse") (Wald, 2002) is a global criteria computed with the standard deviation taken into account the number of bands and the ratio of the PAN and MS resolutions. Lower value indicates a better synthesis of the multispectral quality.

The bias is less than 1 percent for all the bands whatever the method under consideration. The ARSIS MSM3M2 method results in lower standard deviation than the IHS method for the blue band (15 percent versus 24 percent and 25 percent); in similar values for the green and red bands (around 19 percent for the green band and 27 percent for the red band); and in higher values for the NIR band (23 percent versus 14 percent and 15 percent). The IHSW method gives the lowest standard deviations for all bands. The IHSW method also gives the lowest ERGAS values compared to the IHS and ARSIS methods (2 percent versus 5-6 percent).

The variance is negative for the blue band with the IHS method and positive (from 3-4 percent for the NIR band to 11 percent for the Green band) for the three other bands. It is negative for all the bands with the IHSW method (from -4 for the NIR band to -16 for the blue band). The ARSIS synthesized images have positive variances around 10 percent for all bands. We can conclude that all synthesized images resulting from the ARSIS method seems to have the same amount of information loss; that the IHSW synthesized images seems to have some added information, the amount depending of the band under consideration; and that the lost or added information in the IHS synthesized images depends of the band under consideration.

The correlation coefficient is around 0.7 for the blue band and 0.9 for the three other bands for the IHS method. It is around 0.9 for all the bands for the ARSIS method. The IHSW method provides the synthesized images with the highest correlation coefficient (above 0.95 for all bands).

Therefore, from the statistics alone, it can be concluded that the IHSW method provides the best results, *i.e.* preserves more the spectral content of the original MS images, then the ARSIS method and finally the IHS method.

3.2 Assessment of the multispectral quality

We tried to see if the dominant spectra in the original MS images are preserved in the synthesized images. The original MS image was segmented into 20 classes using the Isodata classifier. Only the clusters having over 1000 pixels were considered for the analysis. The first, second and fourth clusters represent tree classes. The third cluster represents the shadow class. The fifth cluster represents the building class. The sixth cluster represents grass, soil and also edge pixels of the vegetation class. The seventh cluster represents roads and

parking. The eighth and ninth clusters represent buildings, parking lots and vacant areas.

The RMSE (Root Mean Square Error) is calculated to help analyzing the spectra of the different clusters. It is given by:

$$RMSE_k = \sqrt{\frac{\sum_{i=1}^{n} MS_k(i) - MS^*_{Syn_k}(i)}{n}} \qquad (1)$$

where n is the number of pixels,

MS_k is the original image for the band k and

$MS^*_{Syn_k}$ is the synthesized image resampled to the original spatial resolution for the band k.
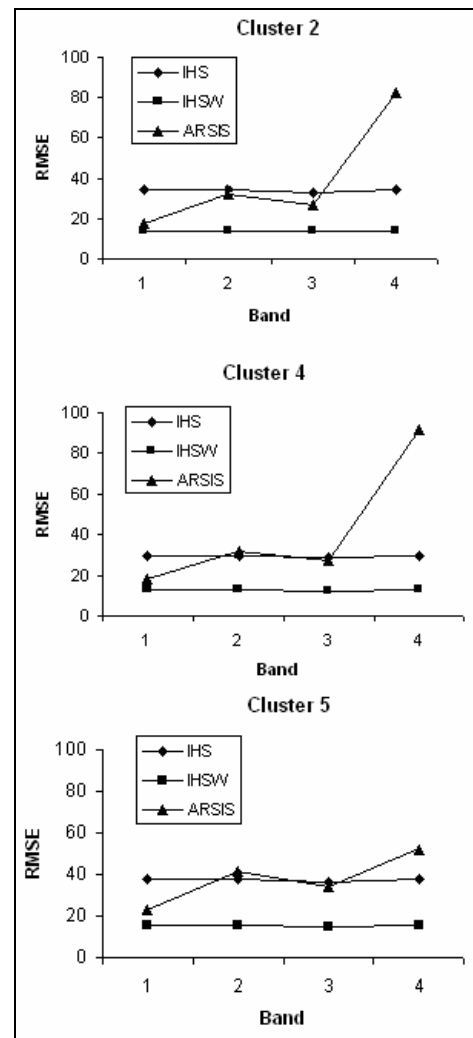

Figure 4. Plots of the RMSE values for cluster 2, 4 and 5.

The RMSE value is calculated for each cluster and each band. Plots of the RMSE values for a few clusters, computed for each method, are given in Figure 4. For the IHS and IHSW methods, the synthesized image of the 421 combination was considered

for the NIR, green and blue bands, and of the 431 combination for the red band.

For each cluster, the RMSE is almost the same in each band for the IHS and the IHSW methods. For example the RMSE for cluster 2 is around 34 in all bands for the IHS method and 14 for the IHSW method. The ARSIS method has different RMSE for each band. The RMSE in the NIR band is high for all the clusters (greater than 45) indicating a higher spectral error in the synthesis of the NIR band compared to the three other bands. Clusters representing vegetation have higher RMSE than the other clusters that represent built-up areas in the NIR band. This may be due to the high brightness values of vegetation in the NIR band.

The analysis of the dominant spectrum of different clusters shows again that the IHSW method provides the best spectral-preserved synthesized images when compared to the ARSIS and IHS method.

4. CONCLUSIONS

The results of the IHS, IHSW and ARSIS methods were compared statistically. The IHSW method provides better statistical results compared to the IHS and ARSIS methods for the case study under consideration. Therefore it is the IHSW method that provides the best spectral-preserved synthesized MS images, followed by the ARSIS method, and then the IHS method. This conclusion is confirmed when analyzing the dominant spectrum of different clusters based on the analysis of the RMSE. However the RMSE are higher for all clusters in the NIR band synthesized by the ARSIS method even if they are lower than the ones resulting from the IHS method for the 3 other bands. It seems that the ARSIS method has some difficulty to really preserve the spectral content of the NIR band

REFERENCES

Couloigner, I., Ranchin, T., 2000. Mapping of Urban Areas: A Multiresolution Modeling Approach for Semi-Automatic Extraction of Streets. *Photogrammteric Engineering & Remote Sensing*, 66(7), pp. 867-874.

Couloigner, I., Ranchin, T., Valtonen, V.P., and Wald, L., 1998. Benefit of the future SPOT-5 and of data fusion to urban roads mapping. *Int. J. Remote Sensing*, 19(8), pp. 1519-1532.

Dutilleux, P., 1987. An implementation of the "algorithme à-trous" to compute the wavelet transform. In *Actes du congrès ondelettes et méthodes temps-fréquence et espace des phases*, Marseille, 14-18 décembre 1987, Springer Verlag Editors, pp.298-304

Núñez, J., Otazu, X., Fors, O., Prades, A., Palà, V., Arbiol, R., 1999. Multiresolution-based image fusion with Additive Wavelet Decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, Vol.37, No.3, May 1999.

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

Wald, L., 2002. *Data Fusion: Definitions and Architectures-Fusion of Images of different Spatial Resolutions*.

Pohl.C, 1997. Tools and Methods used in data fusion. In *Proceedings of the 17th EARSeL symposium on Future Trends in Remote Sensing*, pp.391-399

Csathó, B., Schenk, T., Seo, S., 2003. Spectral Interpretation Based on Multisensor Fusion for Urban Mapping, In: *2nd IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, pp. 8-12. http://www.erc.msstate.edu/~suyoung/publications.htm (accessed 20 Jan. 2005)

ACKNOWLEDGEMENTS