

A METHOD BASED ON LOCAL VARIANCE FOR QUALITY ASSESSMENT OF MULTIREOLUTION IMAGE FUSION

M. Beauchemin^{a,*}, Ko B. Fung^a, X. Geng^b

^a Canada Centre for Remote Sensing, 588 Booth Street, Ottawa, Canada K1A 0Y7 - (Mario.Beauchemin, Ko.Fung)@ccrs.nrcan.gc.ca

^b Prologic Systems Limited, 75 Albert Street, Suite 206, Ottawa, Canada K1P 5E7 – XiaoYuan.Geng@ccrs.nrcan.gc.ca

Commission III, WG III/6

KEY WORDS: Remote Sensing, Image Fusion, Evaluation, Spatial Quality

ABSTRACT:

Several methods exist to combine a panchromatic image of high spatial resolution with lower resolution multispectral imagery. Of particular interest are those methods designed to simulate real multispectral images having the spatial resolution of a panchromatic image. To help justify an algorithm over another one, quantitative evaluation of the quality of a fused image is necessary. In most cases, the evaluation is performed with the original high- and low-spatial resolution images degraded to a coarser resolution by pixel-block averaging. The multispectral image of the highest resolution serves as a reference image (real image). Most approaches proposed for quality assessment are based on statistical measures computed over the whole image; typical measures are the correlation coefficient and the root-mean-square deviation. However, these measures make no reference to the spatial domain. In this paper, we suggest measures based on local variance computed over a three-by-three pixel window as complementary measures to evaluate the quality of the fused images. The rationale is that an ideal fused image must replicate the variance of the reference image when estimated locally. To help discriminate between local variance induced by real details as opposed to artefacts, the variance is partitioned into two terms. Each term takes into consideration the expected direction of the added details over the multispectral image oversampled by pixel replication. The method is illustrated with different fusion models applied to an Ikonos image.

1. INTRODUCTION

1.1 Fusion of Remote Sensing Data

Several methods exist to combine a panchromatic image of high spatial resolution with lower resolution multispectral imagery. Of particular interest are those methods designed to simulate real multispectral images having the spatial resolution of the panchromatic image. Those methods include the Brovey, Intensity-Hue-Saturation, principal components, regression variable substitution and multiresolution approaches (e. g. wavelets, high pass filtering, band correlation). Details on these methods can be found in Pohl and Van Genderen (1998), Ranchin and Wald (2000) and Price (1999).

A restricted number of measures are usually applied to assess the quality of fused images. Most of them are based on statistical parameters such as correlation coefficients and comparison of distribution moments of images (see Wald et al. 1997 and references therein). Because high-spatial multispectral reference images are not readily available, the assessment is usually performed using images obtained by averaging high- and low-resolution images to a lower resolution level. The fusion is then performed between both the spatially degraded multiresolution and panchromatic images. The original multispectral images serves then as reference images for quality assessment.

1.2 Motivation

The statistical measures outlined above are computed over the entire image without considering the spatial feature space. Since these are global statistics, no insight information is gained at the local scale. However, the quality of a fused image at the local

level is an important property that needs to be considered. For example, Hill et al. (1999) stressed the importance of reproducing textural properties for forested applications. In this paper, we present a method for quality assessment based on local variance. In addition to taking into account the spatial context, a way to help differentiate details induced by artefacts from real features is proposed.

Other examples of methods devised to evaluate the spatial quality of image fusion include the high-pass filtered panchromatic image method of Zhou et al. (1998) and the blur parameter estimation method of Li (2000).

1.3 Paper Outline

In the following, because of the nature of the data used, the ratio in pixel size between two successive resolution levels is 1:4. The highest spatial resolution of the original panchromatic image is referred to as L0 (level 0). The next level, L1, corresponds to the highest spatial resolution of the multispectral image. Performance assessment is done at level L1. The ‘synthetic’ image (L1) will be referred to the image resulting from the fusion process. The ‘reference’ image corresponds to the ‘real’ multispectral image (original multispectral image, L1). The ‘replication’ image refers to the multispectral resolution image at L2 up-sampled by replication to equate the resolution corresponding to the reference image (L1).

The paper is organized as follow. The proposed measures are introduced In Section 2. Experiments performed on an Ikonos image are presented in Section 3. Section 4 presents the conclusion.

2. QUALITY ASSESSMENT

2.1 Local Variance

Measures of local variance have been widely used in image processing for texture measures (Gonzalez and Wintz, 1987) and studies of spatial image structure (Woodcock and Strahler, 1987). To measure the local property of the fusion process, we suggest to compute the average value of local variances (ALV) estimated around each pixel of the fused image within a window of 3 by 3 pixels. The ALV computed in this manner shares similarity with semivariance estimated at a lag of one pixel (Jupp et al., 1989).

2.2 Variance Partition

Although two images may have the same global variance (e.g. reference and synthesized), they may nevertheless be poorly correlated. The same statement applies to AVL. Such situations should indicate that the details added over the lower resolution image to simulate the higher resolution image are mainly noise artefacts. This can be expected, for example, when the variance is modelled regardless of the level of correlation. To help distinguish such a case, the variance can be partitioned into two terms.

A diagram illustrating the rationale of the approach is shown in figure 1 for a 1D line profile. The lower multispectral resolution image is first up-sampled by replication. This image represents the 'null' case (no details added). The difference between the reference image and the replication image is then calculated. The same operation is applied between the synthetic image and the replication image. The sign of all pixels in each difference image are next determined. For accurate fusion, one expect that the sign for each pixel of the difference image between the synthetic image and the replication image to be the same as the sign of the pixels of the difference image between the reference image and the replication image. In other words, the details added over the replication image to produce the synthetic image should be in the same direction, compared to the replication image, than the reference image. The variance of the synthesized image can then be broken into two terms, one in which both directions agree, and one in which directions disagrees.

2.3 Quality Assessment

The approach rationale is that an ideal fused image must replicate the variance of the reference image when estimated locally. The following measures were computed: average local variance (AVL), local variance of pixels of same sign (AVLr), local variance of pixels of opposite sign (AVLw) and the ratio of the two aforementioned measures (AVLr / AVLw). In addition to the proposed measures, correlation coefficients and distribution's moments were also computed (global statistics).

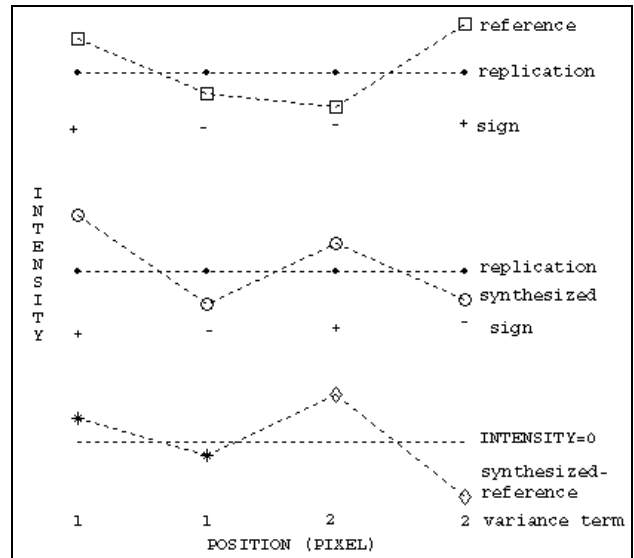


Figure 1. Variance term 1 indicates pixels included in the calculation of variance in which detail direction agrees while term 2 are for disagreement in direction.

3. EXPERIMENTAL RESULTS

3.1 Fusion methods

The ARSIS concept based on wavelet analysis (Ranchin and Wald, 2000) was applied on the same dataset with different model transformations for the wavelet coefficients. Based on their own studies and works from others, Ranchin and Wald concluded that 'the quality of the synthesized images was demonstrated to be the best achievable compared to other fusion methods currently available...'

In the present paper, the adjustment of the wavelet coefficients is computed between 4m (L1) and 16m (L2) pixel sizes (we respect a pixel size ratio of 1:4 as in the original dataset). The transformation methods tested were: A) adjustment for equality of means and variances, and B) adjustment using least-square fit. Details on these models can be found in Ranchin and Wald (2000). Two other models representing slightly modified versions of model B were also tested.

3.2 Data

The data used for testing is a small portion of an Ikonos image. The main land types within this subset include water, forested areas, bogs, unpaved roads, and houses (Fig. 2). The near infrared channel was chosen because, it is the least correlated channel with the panchromatic image ($R^2 = 0.61$) and therefore represents a more challenging case than the three other channels ($R^2 \sim 0.9$).



Figure 2. Original Ikonos panchromatic image subset (1024x1024 pixels, 1m).

3.3 Results

Table 1 provides a summary of the AVL measures applied to the synthesized images (fused images) under different models. Let us consider first methods A and B. One can see that although method A provides the closest ALV value to the reference image (2297 vs. 2514), the ratio ALV_r/ALV_w is higher for method B. This indicates that the proportion of details due to noise is likely higher in method A than in B. Table 2 shows other quality assessment measures applied on the same set of images. The first three rows display the correlation coefficients between wavelets coefficients obtained from L1 to L2 of the reference and the synthesized image. This measure is

very similar in essence to the high-pass filtered panchromatic image method of Zhou et al. (1998). The next rows are respectively the difference in standard deviation between the synthesized image and the reference image, the difference in means, the standard deviation of the difference image between the synthesized image and the reference image, and finally the correlation coefficient between the reference and the synthesized image. Numbers in Table 2 support previous observations based on Table 1 indicating that method B is superior than method A.

Some other experiments were conducted. It appears that better results are achieved when the original panchromatic is degraded by block averaging between L0 and L1, instead of using the low-low image coming from wavelet decomposition. This can be observed in the fourth column of Tables 1 and 2 (method B & aggregation). This result stress the importance of choosing the appropriate filter function to approximate sensor characteristics at different levels of resolution.

A graph of the digital number (DN) of the NIR image as a function of the panchromatic image at L1 is shown in Figure 3. The rather low correlation coefficient ($\rho=0.78$) reflects the observed data scatter. However, the appearance of clusters in Figure 3 is striking and it is obvious that each of these cluster is aligned along different slopes. Image segmentation based on histogram thresholding was applied to the scene. It is based on a pseudo-NDVI index between the NIR channel and the sum of the three remaining ones. This index is computed for each L2 pixel. The three classes obtained from histogram thresholding were identified on the image corresponding to water, vegetation and others. The wavelets coefficients were modelled for each class individually (using model B). Measures for quality assessment are shown in the fifth column of Tables 1 (Method B & aggregation & segmentation). The improvement over method A is noticeable. The ALV_r is equal to 96% the value of method A while ALW is reduced by a factor of 2 (ALV_r / ALW also increase by a factor 2). Most entries in Table 2 also show improvement over all other methods.

Table 1. Local variance measures for synthesized images under different transformation methods.

SYNTHETIC	Method A	Method B	Method B & aggregation	Method B & aggregation & segmentation	replication	reference
Average local variance (ALV)	2297	1779	1851	1965	1369	2514
ALV_r	1830	1468	1625	1747		
ALV_w	453	306	225	218		
ALV_r / ALV_w	4.04	4.80	7.22	8.01		

Table 2. Some statistical measures for synthesized image under different transformation methods.

	Method A	Method B	Method B & aggregation	Method B & aggregation & segmentation	replication
ρ_{hh}	0.779	0.800	0.850	0.861	0.334
ρ_{vv}	0.671	0.730	0.879	0.888	0.326
ρ_{hv}	0.615	0.670	0.769	0.778	0.180
$\delta\sigma$	-1.190	-3.877	-3.485	-2.888	-7.546
$\delta\mu$	-0.504	-0.504	-0.505	-0.507	-0.363
$\delta\sigma_{img}$ diff.	32.524	29.637	23.465	22.363	37.384
ρ_{img}	0.943	0.952	0.970	0.973	0.922

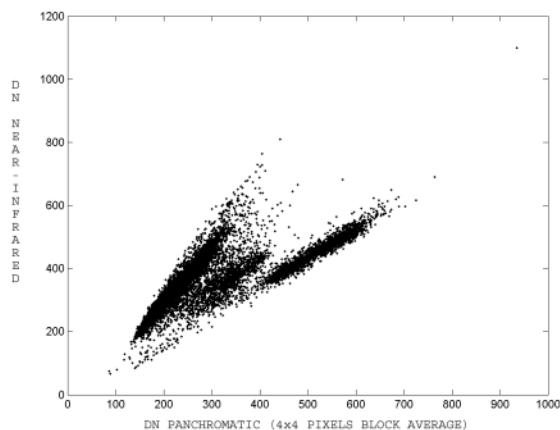


Figure 3. DN values of the NIR image as a function of the panchromatic image DN at level 1 (L1).

4. CONCLUSION

The average value of local variances estimated within a 3 x 3 pixel window was used to gain insight information of images at a local scale. The quality assessment principle relies on the premise that an ideal fused image must replicate the variance of the reference image on a local basis. To help discriminate between local variance induced by real details as opposed to noise artefacts, it was suggested to partition the variance into two terms. The tests conducted in the present study show the sensitivity of the proposed measures to different fusion models in the case of the ARSIS fusion methodology using wavelet analysis. The results (Table 1) support the analysis carried out with some other measures (Table 2).

Regarding the fusion methods themselves, it was demonstrated that transformation models might have to take into account the nature of the land to synthesize images of better quality. Here, improvements for the fused image in the NIR channel were realized by segmenting the scene according to a pseudo-NDVI spectral index.

5. REFERENCES

- Gonzalez, R. C., and Wintz, P., 1987. *Digital Image Processing*, Addison-Wesley, Menlo Park, CA, p. 414.
- Hill, J., Diener, C., Stöver, O., and Udelhoven, Th. 1999. A local correlation approach for the fusion of remote sensing data with different spatial resolutions in forestry applications. In: *International Archives of Photogrammetry and Remote Sensing*, Valladolid, Spain, Vol. 32, Part 7-4-3 W6.
- Li, J., 2000. Spatial quality evaluation of fusion of different resolution images. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Amsterdam, Netherlands, Vol. XXXIII. <http://citeseer.nj.nec.com/403011.html> (accessed 26 June 2002).
- Pohl, C. and van Genderen, J. L., 1998. Multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing*, 19(5), pp.823-854.
- Price, J. C., 1999. Combining multispectral data of differing spatial resolution. *I.E.E.E. Transactions on Geoscience and Remote Sensing*, 37(3), pp.1199-1203.
- Ranchin, T., and Wald, L., 2000. Fusion of high spatial and spectral resolution images: the arsis concept and its implementation. *Photogrammetric Engineering & Remote Sensing*, 66(1), pp. 49-61.
- Woodcock, C. E., and Strahler, A. H., 1987. The factor of scale in remote sensing, *Remote Sensing of Environment*, 21, pp. 311-332.
- 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(4), pp. 743-757.