An Optimal Fusion Approach for Optical and SAR Images

Oguz Gungor and Jie Shan

Geomatics Engineering, School of Civil Engineering, Purdue University
550 Stadium Mall Drive, West Lafayette, IN-47907, USA
Phone: ++1-765-494-2168; Fax: ++1-765-496-1105.
Email: {ogungor, jshan}@ecn.purdue.edu

Commission VII

KEY WORDS: Image Fusion, SAR, Landsat, Ikonos, Quickbird, Multisensor, Spectral, Spatial

ABSTRACT

Remotely sensed images are invaluable to acquire geospatial information about earth surface for the assessment of land resources and environment monitoring. In most cases, the information provided by a single sensor is not complete or sufficient. Therefore, images collected by different sensors are combined to obtain complementary information.

Each remote sensing sensor has its own advantage and disadvantage over the other sensors. SAR (Synthetic Aperture Radar) sensors are active sensors and can collect images during day and night without being affected by weather conditions. SAR sensors are capable of sensing the geometry and structure of the features such as terrain topography, thickness and roughness of surface cover. They also sense the moisture content and presence of vegetation. However, Visible-Infrared (VIR) sensors are passive sensors that sense the electromagnetic energy reflected from surface. Therefore, the information provided by the SAR data alone may not be satisfactory for a detailed analysis of the terrain, since it does not has the capability of collecting spectral information about terrain cover types. For this reason, fusion of VIR and SAR images provides complementary data to increase the amount of information that can be extracted from the individual input images.

For an optimal image fusion, some criteria should be defined for algorithmic development. The success of the fusion strongly depends on the criteria selected. In this work, a pixel based image fusion algorithm is proposed. The new method forms the fused images as the linear combination of the input images. The method employs adaptive windows to establish statistical relationships between the input images to calculate new fused pixels. The fused pixels are calculated using two criteria: 1) Variance of the local window in fused image should be equal to the variance of the corresponding window in higher resolution image to transfer spatial detail. 2) Mean of the local window in the fused image should be equal to the mean of the window in the original lower resolution image to retain the color content.

This paper describes the principles of the proposed approach and assesses its properties. To test the performance of the new approach, the images from different sensors are fused as well as images from same sensors using proposed algorithm and PCA, Brovey and Multiplicative image fusion methods. For the different type of sensors, SAR image is fused with Ikonos, Quickbird (QB) and Landsat Thematic Mapper (TM) XS images. For the same sensor type, QB pan image is fused with QB XS and Ikonos XS images. The results are evaluated visually and analytical derivation and graphic results are presented.

1. INTRODUCTION

Over the last two decades various pixel level (Pohl and van Genderen, 1998) image fusion algorithms have been introduced. Their objective is to obtain a new image that has superior properties over the individual input images with different properties. The input images can be optical images with the same or different spatial resolutions such as Pan or XS images collected by Visible-Infrared (VIR) sensors. They can also be collected by different sensors (e.g., VIR and Synthetic SAR images) at the same or different time.

Expected benefits from the fused images vary, depending on the images used for the fusion. For example, if the anticipated benefit from fusion is to detect the changes occurred in a scene over a period of time, then the images with different acquisition time should be used. The other prevalent application area of image fusion is to enhance spatial resolution of multispectral images using a panchromatic (Pan) image with higher spatial resolution. The main purpose is to get a fused image that retains the spatial resolution from the panchromatic image and color content from the multispectral images. Therefore, a good fusion algorithm should not distort the color content of the original multispectral image while enhancing its spatial resolution.

VIR sensors are passive sensors that sense the electromagnetic energy reflected from surface. Therefore, such sensors offer spectral information about terrain cover types. However, sometimes it is impossible to distinguish some vegetation species using only optical images because of their similar spectral responses. VIR sensors are also not capable of taking images during night and they do not have the ability of imaging targets hidden by clouds, trees, and other ground cover. On the other hand, SAR sensors are active sensors and can collect images during day and night without being affected by any weather condition. SAR sensors are also capable of sensing the moisture content and existence of vegetation, the geometry and structure of the features such as terrain topography, thickness and roughness of surface cover. They also can penetrate materials which are optically opaque, and thus not visible by optical or IR techniques. Low-frequency SAR technology can be used under certain conditions to map the area covered by vegetation by penetrating the foliage (Sandia, 2006). Therefore, SAR images complement photographic and other optical imaging capabilities to increase the amount of information that can be extracted from the individual input images. Therefore fusion of SAR and VIR images are getting popular to increase the amount of information that can be acquired from only VIR or SAR images.

Existing image fusion methods can be categorized into three groups according to their mathematical models. The first group is based on color theory. IHS (Intensity, Hue, and Saturation), Brovey and Multiplicative methods belong to this group. The underlying assumption for the IHS method is that the panchromatic image is equal to the intensity image obtained from the RGB image (Chavez et al, 1991; Liu, 2000). However, this assumption is not always true if the input image has more than three bands or is collected by a different sensor than the panchromatic image. As a consequence, the fusion product will have spectral distortion which causes color deformation in the fused product. Brovey method in principle also uses the intensity calculated from XS bands. This method is based on the chromaticity transform first introduced by (Gillespie et al, 1987). It can be easily shown that Brovey method and Multiplicative assure that the ratio among the original XS bands are kept after fusion. This is an important property to keep the spectral content of the original XS bands after fusion. The success of both methods depends on the intensity calculation from the original multispectral image. All the three methods cause color distortion if the XS bands have a different spectral range than the panchromatic image (Liu and Moore, 1998), which is inevitable if the input images are collected from different sensors.

The second group of image fusion methods employs wavelet transform to construct the fused images. Some wavelet based image fusion methods use the practical implementation of the Mallat algorithm (Li, 1994; Gungor and Shan, 2005a). Others use the “A-trous” algorithm for the wavelet transform and multi resolution analysis (Nunez et al, 1999). Both methods are compared and evaluated in detail in (Gungor and Shan, 2005a). Since only spatial content of the original multispectral image is changed, the spectral content is preserved in the fusion result. As the result, wavelet transform in general has a better performance than the color theory based methods in terms of color conservation. However, it is not as good as the previous methods in terms of spatial enhancement.

The third group is based on statistical properties of the images. As a representative, in the principle component analysis (PCA) approach, the first principle component is replaced by the panchromatic image, and the fused image is obtained through the inverse PCA transform (Zhou et al, 1998). Price (1999) forms the fused image as the linear combination of the input images. The method is successful in preserving the color content of the input XS image; however, produces blocking artifacts (Park and Kang, 2004). This problem is getting worse when different sensors are used. Park and Kang (2004) took Price’s algorithm one step ahead and calculated the linear combination for every single pixel other than for a window. They also developed an algorithm to control the contribution of the high resolution image to the fused image by adaptive gains to incorporate the difference of local spectral characteristics between the high and the low resolution images into the fusion (Park and Kang, 2004).

In this paper, a new image fusion algorithm called σ - μ method (Gungor and Shan, 2005b) is used to fuse SAR and VIR images. σ - μ method forms fused images as the linear combination of the input images. Two criteria are introduced to determine such a relationship based on a moving window computation. The proposed method has an important advantage over the other methods in that the properties of the fusion outcome are known. It is possible to control the amount of spatial gain or spectral content loss in the fused image simply by changing the window size used in fusion process.

This new image fusion approach is tested with two groups of images. First group images are from SAR and VIR sensors. These images include SAR image and QB XS, Ikonos XS and Landsat TM images which have 2.5m and 2.4m, 4m and 30m spatial resolutions, respectively. The other group images come from two different VIR sensors. These images include QB Pan, QB XS and Ikonos XS images that have 0.7m, 2.4m and 4m spatial resolutions, respectively. The fusion of VIR images by σ - μ method is tested and evaluated in (Gungor and Shan, 2005b) visually and quantitatively by comparing its results with the fused images created by PCA, Brovey, Multiplicative image fusion methods. This paper reports the fusion results of SAR and VIR images from σ - μ method for different types of sensors and evaluates the amount of transferred spatial information from SAR image into XS images. Moreover, this paper also includes fusion results of different VIR images and visual comparisons with other methods mentioned above.

2. PRINCIPLES OF THE σ - μ METHOD

The σ - μ method constructs the fused image as the linear combination of the input images

$$F_{k_{(m,n)}} = a_{(m,n)} I_{o_{(m,n)}} + b_{(m,n)} I_{k_{(m,n)}} \tag{1}$$

where m and n are the row and column numbers, k = 1,2, …., N (N = number of multispectral bands); F_k is the fused image, I_0 is Pan image, I_k is the XS band and “a” and “b” are the weighting coefficients for pixel location (m,n). Evidently, the coefficients control the amount of contribution from panchromatic image and multispectral bands respectively. The fused image can be obtained using Equation 1 if “a” and “b” coefficients are determined using the input Pan and XS images.

Rules or criteria must be set to determine the fusion coefficients, which become the key of the image fusion mechanism. The selected criteria determine the properties of the fusion outcome. Considering that the image fusion is to retain the high spatial information or details from the Pan image and spectral information or color from the XS image, the following two criteria are set

1) The local variance in the fused image should be equal to the corresponding local variance in the Pan image, such that its spatial details, described by the variance, can be retained in the fused image. Based on Equation 1 this statement can be expressed as follows

$$Cov(F_i, F_i) = a_i^2\sigma_o^2 + 2a_i b_i \sigma_{oi} + b_i \sigma_i^2 = \sigma_o^2 \tag{2}$$

2) The local mean in the fused image should be equal to the corresponding local mean in the XS image, such that the color content, described by the local mean, is retained in the fused image. Based on Equation 1 this statement can be expressed as follows

$$Mean(F_i) = a_i \mu_o + b_i \mu_i = \mu_i \tag{3}$$

In the above two equations, the notations for window location (m,n) are omitted for a clearer expression. The subscript 0 and i are respectively for the panchromatic image and the i-th band of the XS image.

F_i : is the band *i* of the fused image;

a_i and b_i are the coefficients to be determined to construct the fused pixel;

σ_o^2 and σ_i^2 are the variances of the Pan and XS images respectively,

σ_{oi} is the covariance between Pan image and the i-th band;

Mean (F_i) is the mean of the fused window,

μ_o and μ_i are the mean of Pan image and the i-th band.

The two criteria will yield two equations which in turn allow us to determine the two fusion coefficients “a” and “b”. Combination of Equation (2) and (3) will yield the coefficients a and b. From Equation (3), we obtain

$$a_i = \frac{\mu_i}{\mu_o}(1-b_i) \tag{4}$$

Substituting a_i to Equation (3) will lead to the following 2nd order polynomial equation about the coefficient b_i

$$\left(\frac{\sigma_o^2\mu_i^2}{\mu_o^2}+\sigma_i^2-\frac{2\sigma_{oi}\mu_i}{\mu_o}\right)b_i^2+\left(\frac{2\sigma_{oi}\mu_i}{\mu_o}-\frac{2\sigma_o^2\mu_i^2}{\mu_o^2}\right)b_i + \left(\frac{\sigma_o^2\mu_i^2}{\mu_o^2}-\sigma_o^2\right)=0 \tag{5}$$

Equation (5) has two roots for b_i and hence two corresponding a_i values will be obtained from Equation (4). Because the fusion process is expected to transfer spatial details from the Pan image into the fused one, we keep the solution that “a” is larger than “b”. If “a” is larger than “b” in both root pairs, then the pair that has the largest “a” is picked. If “b” is larger than “a” in both root pair, then the pair that has the largest “a” is picked. The other solution would make more contribution from XS image, which would result in poor spatial detail in the fused image. In some cases, the “b” and “a” coefficients are complex numbers. Under such circumstance, the fusion process essentially can not contribute to the lower resolution image. Under this circumstance, the real components of the complex roots are taken as “a” and “b” coefficients, since they can make Equation (5) closest to zero in the domain of real roots. To determine these coefficients, local windows in both input images will be employed. The “a” and “b” coefficients are determined for the center pixels of the local windows. For implementation details, readers may refer to (Gungor and Shan, 2005b).

3. TESTS AND DISCUSSIONS

This study uses images from four different sensors. Images used include Quickbird panchromatic band (0.7 m) and XS bands (2.7m), Ikonos XS bands (4m), Landsat TM (30m) and SAR image (2.5m) over the Davis-Purdue Agricultural Center (DPAC) area. Image fusion essentially occurs when the involved images have the same spatial resolution. Thus, XS images need to be resampled such that they have the same spatial resolution with the Pan or SAR image. For resampling, the nearest neighbor method is used. Other options are bilinear interpolation, cubic approximation to sinc function and 8-point or 10-point sinc function interpolation methods. These methods interpolate new pixel values using the surrounding neighbor pixels, which changes the spatial distribution and color content of the original image. However, in the nearest neighbor method the new pixel value is assigned as the value of the nearest pixel. i.e., the original pixel values repeat. For this reason, nearest neighbor method is selected as the interpolation method, since the other methods have a deteriorating effect on the original structure of the XS image.

Histogram matching is needed when the input images are collected from different sensors. The images from different sensors will have different brightness levels due to different angles of the sensor platforms and different illumination conditions of the scene resulting from different image acquisition times and the difference between the wavelength extensions of the different satellites. Histogram matching converts the histogram of one image to resemble the histogram of another. Thus, histogram matching has to be applied to the input images such that histogram of the XS image is matched to the histogram of the Pan image. It will adjust brightness level of the XS image with respect to the Pan.

On the other hand, SAR images may have speckle noises because of the nature of the radar imagery. These speckles affect the quality of the fused products since they are transferred into the fused images by image fusion algorithms. For this reason, SAR images should be analyzed before fusion and these speckles should be suppressed using an algorithm. In this study, speckles were smoothed using (Lee, 1981) algorithm.

Original images and their fusion results are shown in Figure 1 and Figure2. Figure 1 contains the fusion results of SAR and VIR sensors, whereas Figure 2 includes the results of VIR sensors. For the same fusion combination, best results from the selected traditional fusion methods, Multiplicative, Brovey and PCA are presented, respectively. All of the fusion results will gain spatial information from the PAN or SAR images. However, the color effect behaves differently from different methods. The fusion of SAR with QB XS, Ikonos XS and Landsat TM images produces interesting and promising results for the σ - μ. As can be seen from the Figure 1, the color of the fused image by the traditional method (PCA and Brovey as an example) is significantly distorted from the original Ikonos XS image. On the other hand, the σ - μ approach reserves the color of the input XS image at the cost of blurring certain fine details. For the fusion of images from the same sensor Quickbird, the first row in Figure 2 suggests that the σ - μ method produces results similar to other traditional methods, however, with slightly less color distortion. For the fusion of different optical sensors (Quickbird Pan with Ikonos XS) in the second row, the traditional method (Brovey as a representative in this case) quite considerably transfers the spatial information from the panchromatic image to the fusion result, such that certain features that exist or are visible only in the panchromatic image become easily visible in the fusion outcome. Such examples include a couple of water bodies in the upper part of the image. They are less visible in the fused image obtained from the proposed σ - μ approach, which yields, as a trade-off, very compatible color with minimum color distortion comparing to the input Ikonos XS image.

One advantage of the proposed method over the other existing fusion methods is that it is possible to control the amount of spatial gain or spectral content loss in the fused image by changing the window size used in fusion process (Gungor and Shan, 2005b). Usually, the larger the window size, the more spatial detail is transferred from the pan image, whereas the more color distortion occurs in the fused outcome. Figure 3 and Table 2 supports this statement. Figure 4 contains the fusion results of the same image with different windows sizes. As seen from Figure 4, when the window size is enlarged, more spatial detail is transferred from pan image, while less spectral content can be kept from original XS image. Table 2 also displays this reality numerically. The correlation coefficients among QB Pan vs. fused XS bands and Ikonos XS vs. fused XS bands are given in Table 2 and plotted in Figure 3. The correlation coefficients among QB Pan and fused XS bands are increasing, indicating that the spatial detail from pan is transferred better when larger window size is used. Conversely, correlation coefficients among original and fused XS bands are decreasing, representing that fused images are sacrificing the color when larger window sizes are used. Therefore, the new approach gives the flexibility to the user that if color is more important than the spatial detail gain in the fused product, then user should use smaller window size. Conversely, if the spatial enhancement is more important than the color information, then larger window sizes should be preferred. However, the determination of an optimal window size, preferably adaptive to the image content, is not trivial. It remains to be an unresolved issue for further investigation. On the other hand, only band-4, which is infrared band, behaves differently such that correlation coefficients among pan and fused XS bands takes small values and coefficients among original and fused XS bands takes negative values. This behavior can be explained by the different nature of the infrared band.

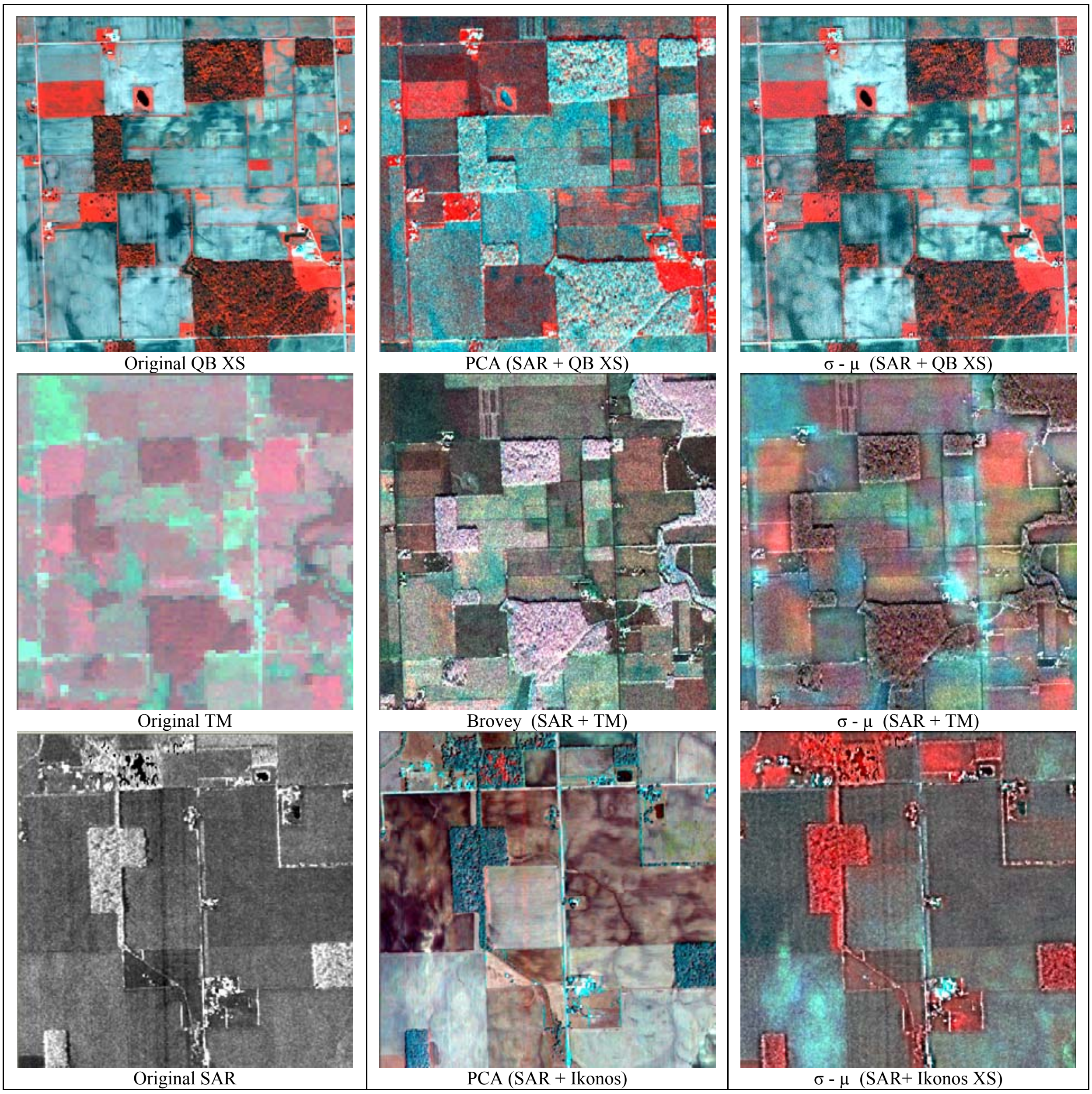

Figure 1. Fusion results of SAR and VIR Sensors. **O**riginal images (left) and fusion results. (middle: conventional methods; right: the proposed method)

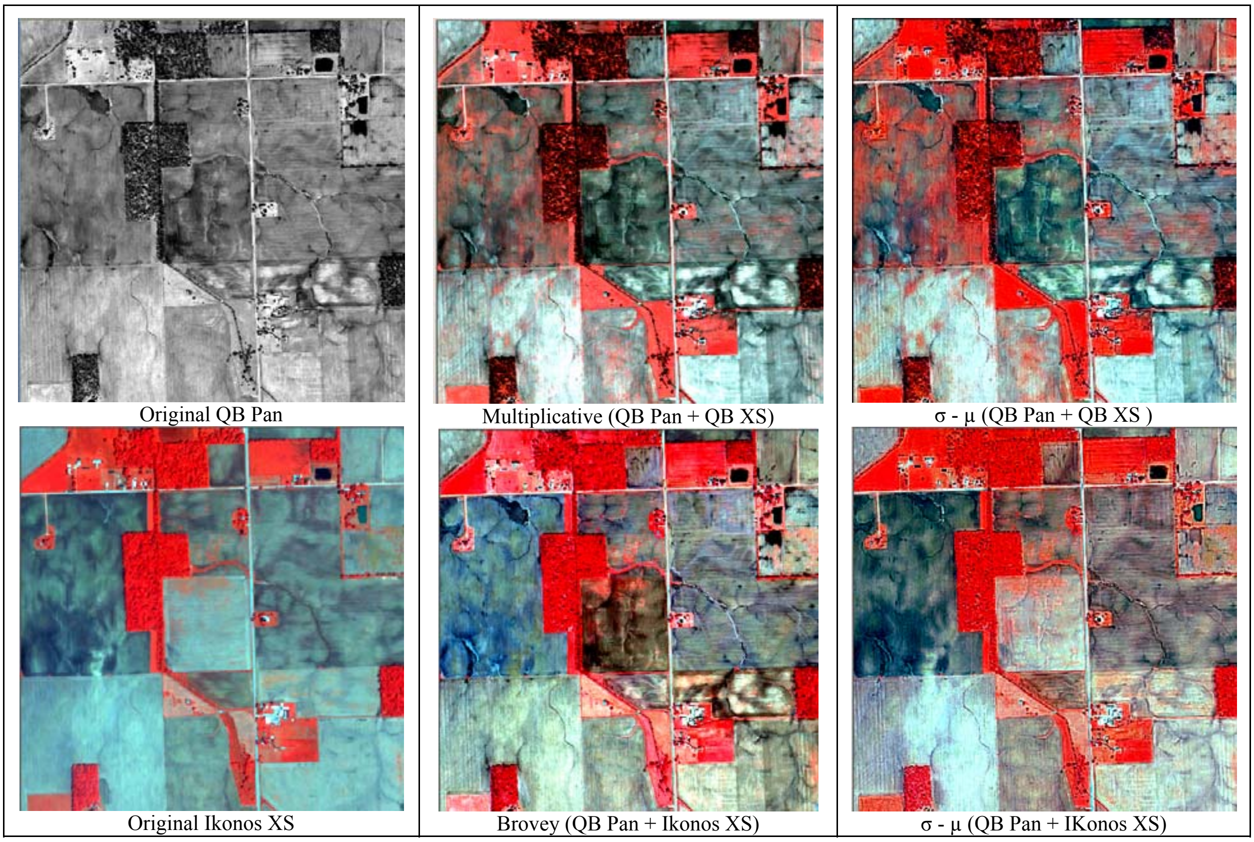

Figure 2. Fusion of VIR Sensors. Original images (left) and fusion results. (middle: conventional methods; right: the proposed method)

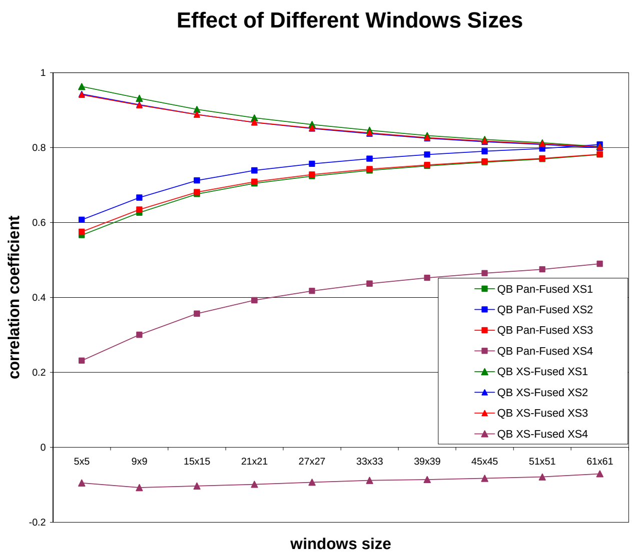

Figure 3. Effect of different window sizes on correlation coefficients

4. CONCLUSION

In summary, the fusion outcome can be regarded as a linear combination of the input images. The properties of the fusion results are characterized by the two fusion criteria. Statistically, the fusion image has the same variance as the input Pan or SAR image and the same mean as the input XS images. The fusion outcome is optimal under these two fusion criteria. Such a modeling allows us to balance the spatial and spectral content of the fusion outcome. It is shown that the fusion window size affects the quality of the fusion results. The window size also affects the computation time. When the windows size is increased, it takes longer time to solve the fusion problem; thus, it takes longer time to produce fused images with the new method. For this reason a window size should be carefully chosen for the optimum solution in terms of spatial detail gain and spectral content loss.

Table 1: Correlation Coefficients among QB Pan vs. Fused XS and Original XS vs. Fused XS bands

	Window Size	**Band 1**	**Band2**	**Band3**	**Band4**
Correlation Coef. among QB Pan and Fused Ikonos XS Bands	**5x5**	0.5665	0.6076	0.5751	0.2316
	9x9	0.6264	0.6665	0.6344	0.3003
	15x15	0.6761	0.7124	0.6813	0.3567
	21x21	0.7044	0.739	0.709	0.3924
	27x27	0.7241	0.7569	0.728	0.4175
	33x33	0.7391	0.7706	0.7426	0.437
	39x39	0.7513	0.7815	0.7539	0.4525
	45x45	0.7611	0.7902	0.7633	0.4648
	51x51	0.7695	0.7976	0.7711	0.4751
	61x61	0.7815	0.8081	0.7824	0.4897
Correlation Coef. among Original and Fused Ikonos XS Bands	**5x5**	0.963	0.9432	0.9411	-0.095
	9x9	0.9316	0.9145	0.9135	-0.1076
	15x15	0.9022	0.8883	0.888	-0.1031
	21x21	0.8794	0.8673	0.8677	-0.0989
	27x27	0.8616	0.8513	0.8524	-0.0934
	33x33	0.8464	0.8375	0.8392	-0.0883
	39x39	0.8322	0.8248	0.8268	-0.0862
	45x45	0.8218	0.8157	0.8179	-0.0829
	51x51	0.8129	0.8081	0.8102	-0.079
	61x61	0.8029	0.7999	0.8018	-0.0707

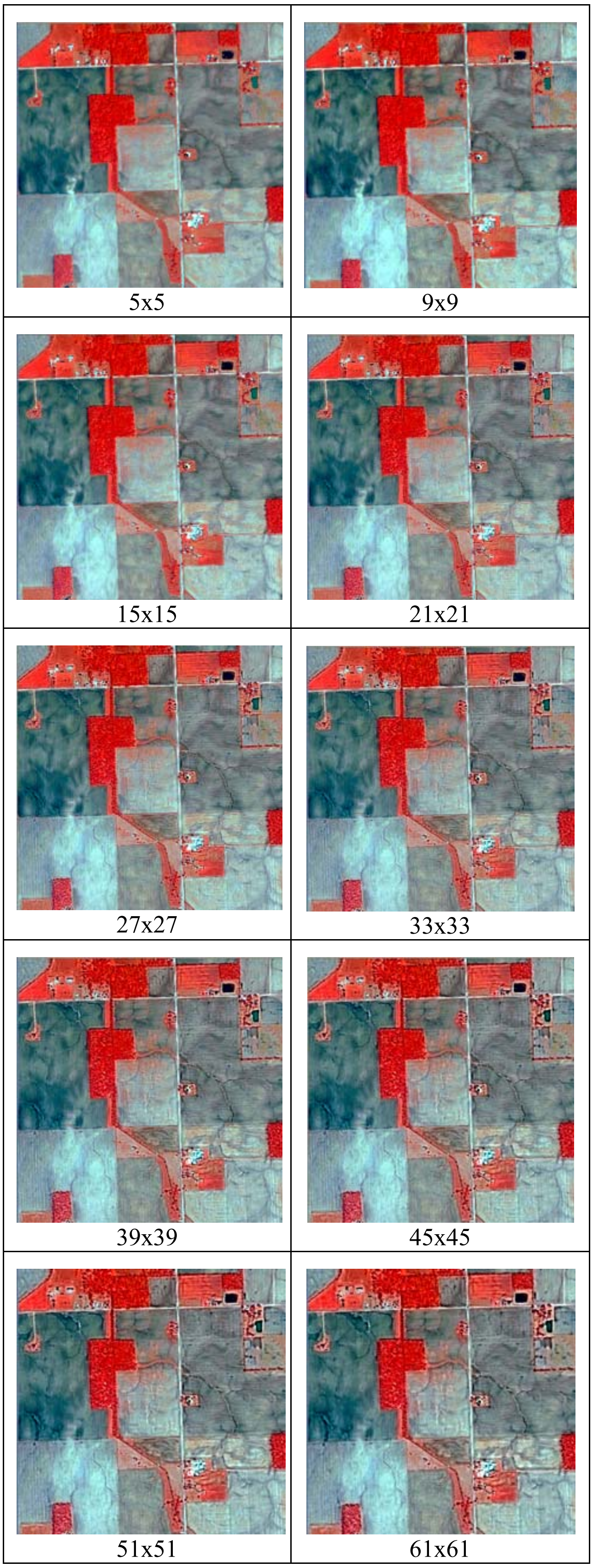

Figure 4. Fusion results with different window sizes (QB Pan and Ikonos XS)

REFERENCES

Chavez, P.S. Jr., Sides, 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*, vol.57 (3), pp.295-303.

Gillespie, A. R. et. al., 1987. Color Enhancement of Highly Correlated Images. II. Channel Ratio and "Chromaticity" Transformation Techniques. *Remote Sensing of Environment*. Vol:22 pp. 343-365.

Gungor, O. and Shan, J., 2005a. A Comparative Study On Wavelet Transform Based Image Fusion Methods Using High Resolution Satellite Images. *ASPRS 2005 Annual Conference "Geospatial Goes Global: From Your Neighborhood to the Whole Planet"* March 7-11, 2005, Baltimore, Maryland

Gungor, O. and Shan, J., 2005b. A Statistical Approach To Multiresolution Image Fusion. *ASPRS Pecora 16 "Global Priorities in Land Remote Sensing"* October 23 – 27, 2005, Sioux Falls, South Dakota.

Lee, J.S., 1981. "Speckle Analysis and Smoothing of Synthetic Aperture Radar Images." *Computer Graphics and Image Processing*. 17(1) pp. 24-32.

Li, H., 1994. Multi-sensor image fusion using the wavelet transform. ICIP-94., *IEEE International Conference*, Vol. 1, 13-16 pp. 51-55.

Liu, J.G. and Moore, J. M., 1998. Pixel block intensity modulation: adding spatial detail to TM band 6 thermal imagery. *Int. J. Remote Sensing*, vol. 19, no. 13, 2477-2491

Liu, J.G., 2000. Smoothing Filter-based Intensity Modulation: a spectral preserve image fusion technique for improving spatial details. *Int. Journal of Remote Sensing*. Vol. 21, No 18, pp.3461-3472.

Nunez , J., X. Otazu, O. Fors, A. Prades, V. Pala, and R. Arbiol., 1999. Multiresolution-based image fusion with additive wavelet decomposition. *IEEE Transactions on Goescience and Remote Sensing*, Vol. 37. No 3.

Park, J. H., Kang, M. G., 2004. Spatially adaptive multi-resolution multispectral image fusion. *International Journal of Remote Sensing*, Vol. 25, No. 23, pp.5491–5508

Pohl, C., van Genderen, J. L., 1998. Reviev Article Multisensor image fusion in remote sensing: concepts, methods and applications. *Int. J. Remote Sensing*, vol.19, No.5, pp. 823-854.

Price, J.C., 1999. Combining Multispectral Data of Differing Spatial Resolution, *IEEE Transactions On Geoscience And Remote Sensing*, Vol. 37, No. 3, May, pp. 1199-1203.

Sandia Laboratories, 2006. *Sandia National Laboratories*. <http://www.sandia.gov/radar/whatis.html> (accessed 15 March, 2006).

Wang, Z.J. et. al,, 2005. A Comparative Analysis of Image Fusion Methods. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 43, No. 6. 1391- 1402

Zhou, J. et. al., 1998. A wavelet transform method to merge Landsat TM and SPOT panchromatic data. *Int. J. Remote Sensing*, vol. 19, no. 4, 743-757