Preprint/Prétirage

Satellite Image Fusion with Multi-scale Wavelet Analysis: Preserving Spatial Information and Minimizing Artifacts (PSIMA)

Yong Du*, Paris W. Vachon, and Joost J. van der Sanden
Canada Centre for Remote Sensing, Natural Resources Canada
588 Booth Street, Ottawa, ON K1A 0Y7
tel: (613)-995-1575
email: Paris.Vachon@ccrs.nrcan.gc.ca

Abstract

Image fusion is the combination of two or more different images to form a new image that contains enhanced information. Consistent with specific application goals, a variety of image products arises from the many available fusion algorithms. However, there is no universal, quantitative performance measure to estimate image fusion quality. The essential objective of image fusion is that nearly all of the original application-specific information should be preserved, and artifacts should be minimized in the final product.

The wavelet transform, a well-known and solid mathematical tool, has already been applied to multi-sensor image fusion. The wavelet transform allows the decomposition of an image into its constituent spatial scale layers. Most image fusion techniques, including wavelet analysis, require that the input images of different spatial resolutions and sample sizes first be re-sampled to achieve spatial registration. The re-sampling could cause a loss of spatial information or might introduce artifacts in the final fused image, especially when the resolutions of the input images are significantly different.

In this paper, as a further development of the application of wavelet analysis to image fusion, we propose a new scheme for multi-resolution image fusion, Preserving Spatial Information and Minimizing Artifacts (PSIMA) with multi-scale wavelet analysis. With the PSIMA scheme, the images are fused in almost their original pixel size. Therefore, the finest spatial information of the input images can be preserved and artifacts minimized in the final fused product. We demonstrate the PSIMA scheme using RADARSAT-1 ScanSAR and NOAA AVHRR images. The results show that the PSIMA scheme is superior to conventional wavelet analysis for image fusion in terms of spatial information preservation and artifact rejection.

Introduction

Remote sensing uses different portions of the electromagnetic spectrum at different spatial, temporal, and spectral resolutions to observe the earth's surface. The multisensor, multitemporal, multiresolution, and multispectral nature of remote sensing data provides countless possibilities for data fusion, which provides an image with more information than any individual image. "Data fusion is a formal framework in which are expressed means and tools for the alliance of data originating from different sources. It aims at obtaining information of greater quality; the exact definition of 'greater information' will depend upon the application." (Wald, 1999). This definition includes all techniques of combining different sources of data to provide more information. In the present paper, we focus only on the fusion of different kinds of satellite image data. Image fusion is the combination of two or more different images to form a new image containing more information by using a certain algorithm (Pohl and van Genderen, 1998).

According to the resulting image products, image fusion can be grouped into two main classes: multi-band color (MBC) and single new image (SNI). The MBC techniques are often carried out during image display or color composition. The components of different bands or channels, in general RGB (red, green, blue), are adjusted by using an algorithm to enhance specific features or to distinguish objects in the image. With the MBC techniques, the original

* Under contract to CCRS from Noetix Research Inc., 265 Carling Ave.,Suite 403, Ottawa, Ontario K1S 2E1 Canada

images may not actually be fused together. The values of an individual band or channel can still be read separately in the constituent display files. On the other hand, the SNI methods generate a single fused image from two or more different images. These two classes of image fusion can be used for different situations. Generally, the MBC technique is for the matching and enhancement of features in the images at the same or similar pixel size, which is relatively straightforward. With the SNI methods, the information from different images can be merged into a single image, which may be useful in overcoming data gaps or as an input physical field with more information, for example. With the MBC techniques, the original images with different pixel sizes must first be re-sampled or interpolated onto the same image grid. In fact, the same first step is usually required with the SNI methods (Li *et al*., 1993; Yocky, 1996, Schowengerdt, 1997; Pohl and van Genderen, 1998). When the pixel sizes of the input images are quite similar, the re-sampling should not cause significant problems. However, if the pixel sizes of the original input images are significantly different, another scheme is required to avoid problems associated with re-sampling.

It is well known that re-sampling the image with the finer pixel grid to match that of the coarser one will result in a loss of spatial information. However, if the image with the coarser pixel grid is interpolated to match that of the finer one, other problems could arise. First, the registration error could be large (on the order of the coarse grid pixel size). Second, the interpolated image, according to the sampling theorem, does not contain any information with frequency higher than that of the original image. Therefore, the interpolated image does not provide any useful information at the finer pixel size. Furthermore, if the interpolated image were decomposed using a multi-scale wavelet analysis, it might appear that fine scale artifact information due to interpolation is present. Therefore, when such decomposition is used in an image fusion operation, some high frequency errors may be introduced.

Although there is no universal, quantitative measurements for image fusion quality, the final product should provide reliable information and be free of artifacts. Multi-scale wavelet analysis has already been used for image fusion, and various algorithms for radiometric normalization have been suggested, but the evaluation for spatial quality of a fused image is still based on visual examination. (Li *et al*., 1993; Yocky, 1996; Zhou *et al*.,1998).

In this paper, we present an image fusion scheme that is based on wavelet multi-scale analysis, Preserving Spatial Information and Minimizing Artifacts (PSIMA) in the final image product. With the PSIMA scheme, the spatial registration is carried out within the closest common scale layer of the two images, after decomposition of the images with a wavelet transform. The wavelet coefficients for common scales from the two images are then fused, and the final image is reconstructed by using an inverse wavelet transform. The original images with their different pixel sizes do not need to be re-sampled onto the same grid, so their original pixel sizes are preserved. The information loss and artifact creation, therefore, are reduced. We demonstrate the PSIMA scheme for the fusion of RADARSAT-1 ScanSAR and NOAA-AVHRR images which have significantly different pixel sizes. Data acquired off the British Columbia coastal area in 1999 (see Fig. 1) were selected (see van der Sanden *et al*., 2000).

Figure 1. *Map of the experimental area off the West Coast of BC.*

Principle of multi-resolution image fusion with wavelet analysis

1. Re-sampling

Assume that we wish to fuse images A and B which have pixel sizes P_a and P_b ($P_a >> P_b$), respectively. Let $P_a = S \times P_b$, and take $S_n \leq S$, to be a power of 2, with *n* chosen to satisfy *min*($S-S_n$). Image A is first registered to image A' with $P_{a}' = S_n \times P_b$, such that the pixel size of image A decreases to the minimum extent possible. If $S = S_n$, the original pixel size of image A will be preserved exactly. The re-sampling in conventional image fusion could be considered to be the special case of *n = 0,* which forces the original pixel size of both images to be re-sampled to a common grid (Yocky, 1996).

2. Decomposition, registration, and resizing

The pixel size of image A□ is now 2^n times that of image B. Image B can be decomposed by using a wavelet transform. The different levels of approximation can be obtained with a pixel size of 2^n times that of the original. For example, if the pixel spacing of image B is 100m, then the pixel spacing of level 1 is $2^1\times100 = 200$m, the pixel spacing of level 2 is $2^2\times100 = 400$m, and that of level 3 is $2^3\times100 = 800$m. Image A□ can be registered to the same grid as image B in its level *n* approximation *(i.e.*, with pixel spacing that is the same as that of image A□). Generally, image B with its original pixel spacing is registered to a certain grid (*e.g.*, a map projection). The registered image B□ is then decomposed by using wavelet transforms to the *n*th level of approximation with the same pixel spacing as image A□. Image A□ can be registered and resized to the *n*th level of approximation of image B□ and retains its pixel spacing.

3. Normalization

When the spectral properties of two input images are different, the format and scale of the digital values in the two files might be very different as well. The comparison of two sets of data requires that they be transformed to the same format, and normalized to a common scale. In MBC techniques, adjusting the histograms of the RGB bands is often used to offset the difference of digital value scales. Therefore, radiometric normalization can be automatically in the display process. For the SNI methods, the digital values from different images are fused together to produce one new image. Generally, important features in the different images should appear in the fused image. Therefore, radiometric normalization may be necessary. One approach is to use a linear regression across the area of interest in the two images. In this case, the values of the digital numbers of different images are adjusted to a common level by using a gain and an offset. Another approach is adjustment of the maximum and minimum values of the two images. Principal Component Analysis can also be used to normalize the input images (Garguet-Duport *et al.*, 1996).

4. Fusion of wavelet coefficients

In wavelet-based image fusion, usually the two images are re-sampled to the same pixel spacing prior to wavelet decomposition (see Fig. 2) (Li *et al.*, 1993; Yocky, 1996). The wavelet coefficients at the various scales are then analyzed in the corresponding layers. If the pixel sizes of the input images are on a similar scale, the re-sampling may not result in a significant error. Therefore, conventional multi-scale wavelet analysis worked well for this kind of image fusion. In practice, however, the pixel sizes of the input images may be significantly different. In our case, the quasi-original image A□ is the same size and has the same pixel size as the *n*th level decomposition of image B□ (see Fig.3). It should be noted that B□ records feature with scales from $P_{b\square}$ to $P_{a\square}$ and greater than $P_{a\square}$, but A□ only records features with scales greater than $P_{a\square}$. That is, there are no features with scale smaller than $P_{a\square}$ in image A□. Therefore, the wavelet coefficients of A□ and B□ may be fused from the *n*th level of the B□ decomposition, which corresponds to the zero level of A□. Depending on the application, there are several different image fusion algorithms. One involves selecting the average value of two sets of wavelet coefficients, and another involves selecting the maximum of the two sets as a new set (Li *et al.*, 1993). A third approach uses Principal Component Analysis to choose the principal component as the new set of coefficients. Note that wavelet coefficient fusion can also be carried out at higher levels by using other common pixel spacings and sizes of A□ and B_n□.

5. Reconstruction

The wavelet coefficients of two images can be calculated and combined into a single new set, which can be considered as a decomposition of a new image. By using the inverse wavelet transform, the new set of wavelet coefficients can be used to reconstruct the fused image. The new image contains some features with scale greater than $P_{a\square}$ from A□ and all features with scale greater than $P_{b\square}$, and smaller than $P_{a\square}$, and some features with scale greater than $P_{a\square}$, from B□. This reconstructed image rejects the features smaller than $P_{a\square}$ from A□, which is an important factor in avoiding interpolation-induced artifacts. The original pixel size $P_{b\square}$ is maintained in the reconstructed image without any loss of spatial information.

Demonstration of PSIMA Scheme

In this demonstration, RADARSAT-1 data were acquired in ScanSAR Wide Beam mode with a 450km swath and 50m by 50m pixel size. The image was acquired on October 15, 1999 at 02:12 (UTC). A NOAA AVHRR image was acquired about 10 hours earlier on October 15, 1999 at 16:13 (UTC) with 3000km swath, 1000m by 1000m pixel size, and 5 channels. In this case, only the AVHRR channel 4 (thermal) is considered. It was necessary to select data sets with 10 hours difference in acquisition time in order to find an AVHRR image that was cloud free. The SAR image was first averaged to 100m by 100m pixel size to reduce speckle noise. The AVHRR image was then interpolated to an 800m by 800m pixel size and registered to the SAR image, following the re-sampling rule of the PSIMA scheme. In this example, $P_{a\square}$ = 800m and $P_{b\square}$ =100m. A common sub-area was selected with both SAR (2048×2048 pixels) (see Fig. 4***a***) and AVHRR (256×256 pixels) (see Fig. 4***b***) data. This pre-processed image pair was then fused using four different image fusion schemes. For a fair comparison, bilinear interpolation was used to resample all images and subsequent image smoothing operations were not used.

Case 1. PSIMA Scheme

Using the PSIMA scheme, the SAR image was decomposed with ***db4*** wavelets (Daubechies, 1988) to three levels: 1024×1024, 512×512, 256×256. The AVHRR image was kept in its original size 256×256 (level zero). The original AVHRR image (256×256) was then registered to the level 3 approximation of the decomposed SAR image (256×256). Because of the different scales of DNs between the SAR and the AVHRR, the wavelet coefficients of the two images were normalized prior to the fusion calculation. By using a linear polynomial fit, a combination algorithm of maximum and average was used. Comparing the absolute values of the wavelet coefficients from both images, the fusion of the wavelet coefficient was constructed with the weighted average of the two values. The wavelet coefficients with the higher absolute value gets weight 0.7, otherwise, the weight is 0.3. In the detail images of the first three levels, only the wavelet coefficients of the SAR are considered because the AVHRR image does not contain spatial information at scales that are less than 800m. The inverse wavelet transform was used to reconstruct the final fused image (Fig. 4***c***) with the fused wavelet coefficients, including levels 1, 2 and 3 details, and the level 3 approximation image. The size of the final fused image is 2048×2048 pixels with a 100m by 100m pixel size. Following the processing, there should be no residual artifacts, and the fine pixel size of the SAR image is maintained.

Case 2. Re-sampling finer pixel size to coarser one

In this case, the SAR image was re-sampled to an 800m by 800m pixel size with a size of 256×256 pixels. Both the re-sampled SAR image and the AVHRR image were then decomposed to two levels: 128×128 and 64×64. The wavelet coefficients of both sizes were normalized and the maximum selection rule was used to fuse the two levels, respectively. The final fused image (Fig. 4***d***) was then reconstructed with the composed wavelet coefficients. In this case, there should be no artifacts produced, but the pixel size of the fused product is 800m by 800m. Therefore, spatial information was lost through this process.

Case 3. Re-sampling both images onto a common grid

In this case, both the SAR and the AVHRR images were first re-sampled to the same grid with an intermediate pixel size of 200m by 200m and a size of 1024×1024 pixels. After normalization, the digital number values of the two images were added pixel by pixel, and divided by 2. The final fused image (Fig. 4***e***) is obtained directly.

Case 4. Over-sampling coarser pixel size to finer one

In this case, the pixel size of the AVHRR image was over-sampled with interpolation to 100m by 100m from 800m by 800m and the size dilated to 2048×2048 pixels from 256×256 pixels. Both images were then decomposed to three levels: 1024×1024, 512×512 and 256×256. Linear polynomial fitting was used to normalize the wavelet coefficients of the two images. The combination algorithm was also used to select the composed wavelet coefficients from all details and approximation. Obviously, the creation of fused wavelet coefficients in Case 4 is different from the PSIMA scheme. Because the interpolated AVHRR image was also decomposed into three levels, the whole wavelet coefficients, including approximation and details of three levels of two images, were used to construct the composed wavelet coefficients. The final fused image (Fig. 4***f***) was then reconstructed using the fused wavelet coefficients.

Discussion and comparison

The fusion of RADARSAT and AVHRR images is a challenge because of the following differences:

- Significantly different pixel size (100m by 100m vs. 1000m by 1000m)
- Use of different portions of the electromagnetic spectrum (microwave vs. infrared and optical)
- Different sensors (SAR, Scanning Radiometer)
- Different observation type (active vs. passive)

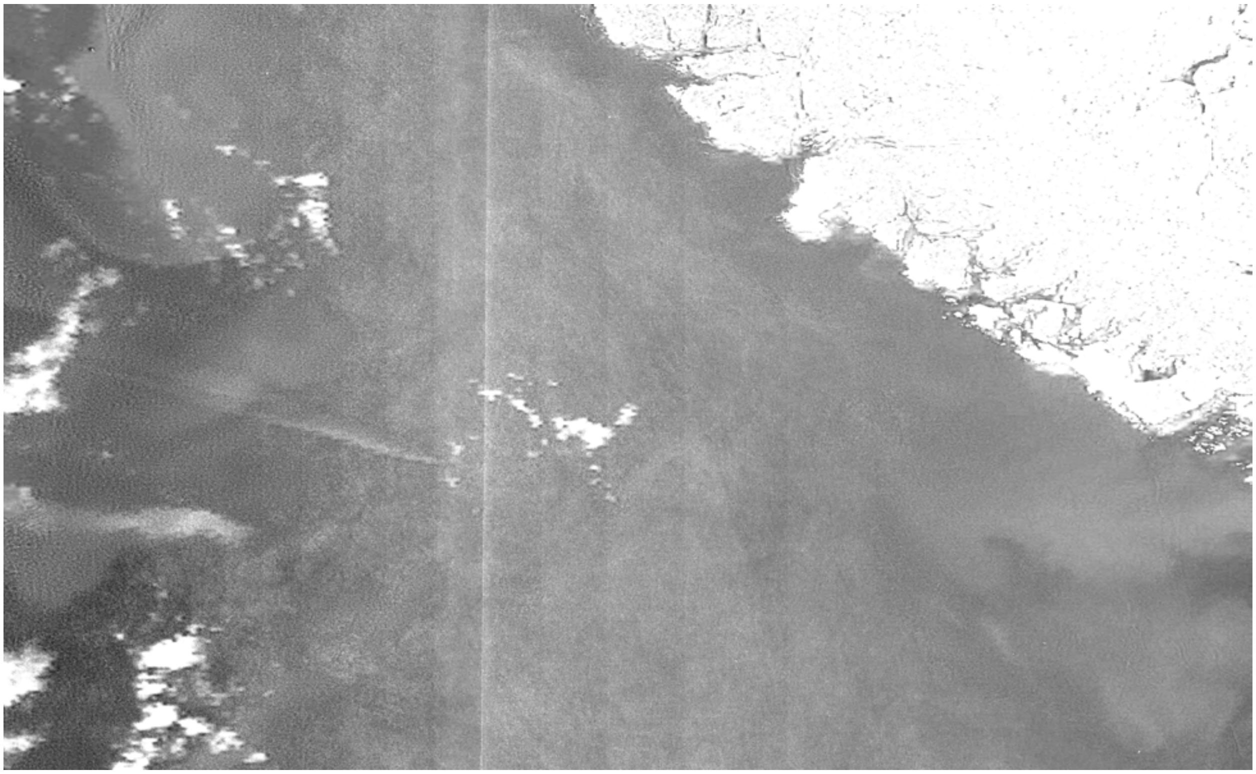

Figure 5. *The image fusion result for the over-sampling scheme with maximum selection rule.*

Therefore, different types of information from the earth's surface are detected by RADARSAT and AVHRR. This kind of image fusion should provide more information than any individual image by retaining the unique features of both images in a single scene. The features from different images cannot be substituted for one another because they contain inherently different information. The objective of this exercise is to match the large-scale features in the two images to explore the relationship between thermal or optical properties and roughness on the sea surface. Generally, the composition of the wavelet coefficients using the maximum selection rule tends to enhance strong features and reject the weaker ones. The average algorithm used for the overlay of images retains all features, but with a contrast degradation.

By comparing the fused images from the four cases considered, we can see that the PSIMA scheme provides a high quality fused image with the original finer pixel size and the SAR images preserved (see Fig. 4). In comparison with Cases 2 and 3, the pixel sizes are degraded to 800×800m and 200×200m, respectively, which results in a lose of spatial information from the SAR (Figs. 4***d*** and 4***e***). As such, the PSIMA scheme is superior to Cases 2 and 3. Our subsequent discussion focuses on the comparison between Case1 (PSIMA) and Case 4.

Because of the resolution of the printer and different enhancements, the differences between Fig. 4***c*** and Fig. 4***f*** are not obvious through visual inspection. A quantitative comparison is necessary. For the fused results of these two schemes and the input images, the average power of the wavelet coefficients in different decomposition scales were calculated (Table 1). The average power of the wavelet coefficients is similar to the power spectrum of the images. Here, the different spatial scales correspond to the different spatial frequencies. It can be seen that the powers of all detail images of the fusion product with the PSIMA scheme, including both maximum and combination algorithms, are exactly the same as that of the RADARSAT data. There is not any other high frequency component involved. Only the lower frequency component (the level 3 approximation) is fused.

For Case 4, after interpolation, the average power of the approximation increases more than 60 fold. At the same time, minor power changes occur in the details of other levels. When they are normalized as a whole, the power in the details is suppressed. With the maximum algorithm, the fusion result of over-sampling is almost the same as that of the PSIMA scheme.

Table 1. Power distribution of wavelet coefficients at different sacles with different schemes										
	RADARSAT	**AVHRR**	**PSIMA**				**Over-sampling**			
			AVHRR		**Fusion**		**AVHRR**		**Fusion**	
			Original	Normalized	Maximum	Combination	Original	Normalized	Maximum	Combination
Level1 Detail	29.34	0.00	0.00	0.00	29.34	29.34	0.01	0.01	29.34	14.38
Level2 Detail	111.51	0.00	0.00	0.00	111.51	111.51	0.17	0.04	111.51	54.64
Level3 Detail	489.91	0.00	0.00	0.00	489.91	489.91	9.47	1.83	490.53	240.41
Level3 Approximation	6.19E+04	4.43E+03	4.43E+03	5.44E+04	7.33E+04	6.26E+00	2.84E+05	5.44E+04	7.31E+04	6.25E+04
Normalization of AVHRR by using linear fitting:										
			PSIMA				**Over-sampling**			
			AVHRRnor= **4.17** *AVHRRori -47.73				AVHRRnor= **0.44** *AVHRRori- **0.10**			

This experiment shows that the fusion results with maximum algorithm, for both the PSIMA and over-sampling schemes, are not satisfactory since some features of the RADARSAT image are lost (Fig. 5). In the fusion result with the combination algorithm and the over-sampling scheme, the power of all details decreases about 50%.

For further exploration, image profiles were

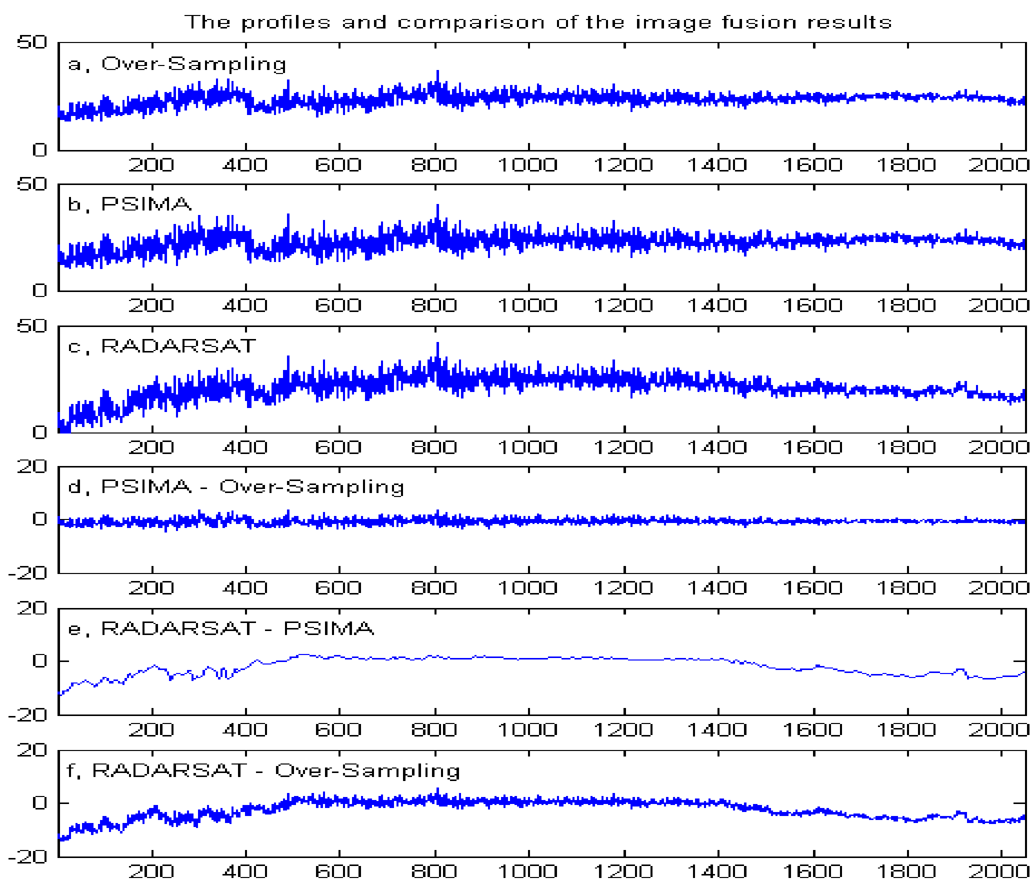

Figure 6. *Image profiles and comparisons of the image fusion results from the different schemes with the combination algorithm (x-axis: distance in pixels; y-axis: digital value)*
a) Case 4, Over-sampling scheme; b) Case 1, PSIMA scheme; c) The original RADARSAT image; d) Difference between Case 1 and Case 4; e) Difference between RADARSAT and Case 1; f) Difference between RADARSAT and Case 4.

selected from the RADARSAT image and the fusion results of both Cases 1 and 4 with the combination algorithm (Fig. 6). It is evident that the difference between the RADARSAT image and the result of the over-sampling scheme includes both high and low frequency components (Fig. 6***f***). The difference between the RADARSAT image and the result of the PSIMA scheme only has low frequency components (Fig. 6***e***), which are from the AVHRR image. The difference between the results of the PSIMA scheme and the over-sampling scheme (Fig. 6***d***) only has high frequency components. From Fig. 6d, we know that some of the high frequency components are lost (difference > 0) and some artifacts are introduced (difference < 0) with the over-sampling scheme.

From the above analysis, it can be seen that previous fusion algorithms introduce errors as long as the SAR's detail image is at a finer scale than the quasi-original pixel size of the AVHRR image. The PSIMA scheme can avoid this kind of error completely.

It has been pointed that the electromagnetic properties of the RADARSAT and AVHRR images are different. The radiometric normalization (or consecutive radiometry) is also an important criterion of image fusion with SNI methods, which determines the selection of image data for fusion depending on the marine phenomena of interest. Because of the objective of this paper, the radiometric normalization was not studied in detail. Both RADARSAT and AVHRR images were treated as digital data with different resolutions, so the conclusion of this paper can be applied to any kind of image fusion with SNI methods.

Conclusions

The PSIMA scheme provides a rigorous method for multi-resolution image fusion by using wavelet analysis, and constitutes an improvement over existing wavelet-based image fusion techniques. We recommend that the PSIMA scheme should be used for image fusion when there is a significant difference in pixel sizes between the original input images. When the pixel size difference is a factor 2 or more, then the PSIMA scheme is appropriate.

It should be noted that even with the PSIMA scheme, the pixel size could not be completely maintained in some cases. The original pixel sizes must be adjusted such that there is a power of two relationship between them. This step might introduce some artifacts. However, the PSIMA scheme is a useful advance, especially when the pixel sizes are significantly different.

Future research will include the application of the PSIMA scheme to multi-sensor and multi-resolution image fusion for marine remote sensing. The radiometric normalization between the RADARSAT and AVHRR images should be considered in detail for specific applications.

References

Daubechies, I. 1988, Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*. Vol. 41: 909-996.

Garguet-Duport, B., J. Girel, J.-M. Chassery, and G. Pautou, 1996, The use of multiresolution analysis and wavelet transform for merging SPOT panchromatic and multispectral image data, *Photogrammetric Engineering & Remote Sensing*, Vol. 62, No. 9: 1057-1066.

Li, H., B. S. Manjunath, and S. K. Mitra, 1993, Multisensor image fusion using the wavelet transform, *Comput. Vis., Image Process: Graph. Models Image Process*., Vol. 57: 235-245.

Pohl, C., and J. L. van Genderen, 1998, Multisensor image fusion in remote sensing: Concepts, methods and applications*, International Journal of Remote Sensing*, Vol. 19, No. 5: 823-854.

Schowengerdt, R. A., 1997, *Remote Sensing, Models and methods for image processing*, Academic Press: 378-387.

van der Sanden, J. J., P. W. Vachon, and J. F. R. Gower, 2000, Combining optical and radar satellite image data for surveillance of coastal waters. *Proceedings of the Sixth International Conference on Remote Sensing for Marine and Coastal Environments*, May 1-3 2000, Charleston, Vol I, pp109-116.

Wald, L., 1999, Some term of reference in data fusion, *IEEE Transactions on Geoscience and Remote Sensin*g, Vol. 37, No. 3: 1190-1193.

Yocky, D. A., 1996, Multiresolution wavelet decomposition image merger of Landsat Thematic Mapper and SPOT panchromatic data, *Photogrammetric Engineering & Remote Sensing*, Vol. 62, No.9: 1067-1074.

Zhou, J., D. L. Civco and J. A. Silander, 1998, A wavelet transform method to merge Landsat TM and SPOT panchromatic data, *International. J. Remote Sensing*, Vol.19, No. 4, 743-757.