Comment. Liu ”smoothing filter based intensity modulation: a spectral preserve image fusion technique for improving spatial details”.

Lucien Wald, Thierry Ranchin

Wald L. and Ranchin, T., 2002. Comments on the paper by Liu "smoothing filter based intensity modulation: a spectral preserve image fusion technique for improving spatial details". International Journal of Remote Sensing, 23(3), 593-597.

COMMENT
LIU "SMOOTHING FILTER-BASED INTENSITY MODULATION: A SPECTRAL PRESERVE IMAGE FUSION TECHNIQUE FOR IMPROVING SPATIAL DETAILS"

Lucien Wald and Thierry Ranchin

Groupe Télédétection & Modélisation, Ecole des Mines de Paris,
BP 207, 06904 Sophia Antipolis cedex, France.
Phone: +33 (0) 4 93 95 74 49 - Fax: +33 (0) 4 93 95 75 35
E-mail: lucien.wald@cenerg.cma.fr

Short title: assessment of quality of methods for the fusion of images

Abstract

We discuss a critical point of the paper. It is shown on the one hand that the author of this paper has mistaken the change of spectral content with the change in spatial resolution, and on the other hand, that the protocol used to establish the advantages of his smoothing filter-based intensity modulation (SFIM) technique over other methods is not appropriate at all. Though the SFIM technique has its merits, the above-mentionned article does not demonstrate its expected qualities.

1. Introduction

Liu (2000) recently published a paper entitled "Smoothing filter-based intensity modulation: a spectral preserve image fusion technique for improving spatial details". This paper presents a technique called smoothing filter-based intensity modulation (SFIM) that permits the fusion of images having a given spatial resolution with images having a higher spatial resolution. The SFIM aims at producing synthetic images (fused products) having the highest spatial resolution in the same multispectral bands of the original low resolution images. The example given in the paper is the fusion of a Landsat TM multispectral image with a resolution of 30 m with a SPOT panchromatic image with a resolution of 10 m. The result is supposed to be a multispectral image with the same bands than the Landsat original image but with a spatial resolution close to 10 m.

The SFIM technique is based upon the extraction by a filtering technique of the high frequencies of the SPOT image and their injection into the Landsat imagery. The SFIM technique is a refinement of the method of Pradines (1986) and of that of Liu and Moore (1998, improperly referred to as Guo and Moore in Ranchin, Wald 2000a). All belong to the group of methods known as the ARSIS concept (Ranchin, Wald, 2000a).

Liu indicates that the IHS and Brovey techniques cause spectral distortion and that "preserving the original spectral properties is very important for most remote sensing applications". Liu makes efforts to establish the advantages of the SFIM technique. The colour composites are convincingly demonstrating that point. However we cannot agree with the following content of his article where he pretends to analyse better the spectral preservation property of the SFIM technique based on statistics and demonstrate the superiority of the SFIM technique over the IHS and Brovey techniques when spectral content is at stake.

The first point is that many articles have been published that demonstrate that the spectral content of an image changes as the spatial resolution changes. Hence it cannot be demonstrated that the fused multispectral images should have the same spectral content than the original Landsat images (preservation of the original spectral properties). Ranchin, Wald (2000a) note that several authors dealing with the methods under concern (including ourselves) improperly write "preservation of spectral content", an inappropriate shortcoming for "high-quality transformation of the multispectral content when increasing the spatial resolution". But here, we do not believe that this is only a problem of expression because the protocol used by Liu for assessing the quality of the transformation is very inappropriate.

2. The three properties of the fused images

Let denote the acquired images of lowest spatial resolution by B_l, and the images of highest spatial resolution by A_h. The subscripts l and h denote the spatial resolution of images B or A, *i.e.*, low and high resolution, respectively. $B^{interp}{}_h$ denotes the result of the interpolation (resampling) of B_l from resolution l to h. Here the images B_l are the Landsat images of resolution 30 m, the $B^{interp}{}_h$ the Landsat images resampled to 10 m superimposable to the image SPOT panchromatic (image A_h) of resolution 10 m. Within the set of images B, B_{kl} denotes the image acquired in the spectral band k. The fusion methods aim at constructing synthetic images $B^*{}_h$, which are close to reality. The methods should perform a high-quality transformation of the multispectral content of B_l, when increasing the spatial resolution from l to h. Wald *et al.* (1997) establish that these synthetic images B^* must respect the three following properties.

First property. Any synthetic image $B^*{}_h$ once degraded to its original resolution l, should be as identical as possible to the original image B_l. Approximation induced by the resampling of $B^*{}_{kh}$ into $B^*{}_{kl}$ should be taken into account: typically the root of the mean of the squared differences $(B_{kl} - B^*{}_{kl})$ on a pixel basis should not exceed 0.05 times the mean value of B_{kl}.

Second property. Any synthetic image $B^*{}_h$ should be as identical as possible to the image B_h that the corresponding sensor would observe with the highest spatial resolution h, if existent. The second property does not imply an accurate synthesis of the multispectral properties of the set B when increasing the spatial resolution. This should be an additional property.

Third property. The multispectral set of synthetic images $B^*{}_h$ should be as identical as possible to the multispectral set of images B_h that the corresponding sensor would observe with the highest spatial resolution h, if existent.

Liu does not check any of these properties.

3. The protocol for quality assessment

Only one protocol has been proposed up to now, which permits to check these three properties in case the images B_h are not available, like in the case of the article of Liu. The first mention of such a protocol originates from Munechika *et al.* (1993), Mangolini *et al.* (1995) extended it. It has been strongly refined and justified by Wald *et al.* (1997). Vrabel (2000) presents a detailed protocol in use in the US Defence for the assessment of fused products for visual analysis. Wald (2001) integrates both protocols in a general frame. It permits to alleviate the need for a reference image if not available and offers a complete checking of the three properties.

In this particular case, the protocol is as follows:

- the fusion method is applied to the original sets of images A_h and B_{kl}. It results into a new set of synthesized images B^*_{kh} at resolution h,
- testing the first property: *any synthetic image B^*_{kh}, once degraded to its original resolution l, should be as identical as possible to the original image B^*_{kh}*. To achieve this, the synthetic image B^*_{kh} is spatially degraded to an approximate solution $(B^*_{kh})_l$ of B_{kl}. If the first property is true, then $(B^*_{kh})_l$ is very close to B_{kl}. The difference between both images is computed on a per-pixel basis. The fused products together with the difference images are visually compared to the original images B_{kl} in order to detect trends of error, if any, possibly related to the objects in the scene. Then some statistical quantities are computed to quantitatively express the similarities and discrepancies between both images.
- a change of scale is performed for the second and third properties. Two sets of images A_l and B_{kv} are created from the original sets of images A_h and B_{kl}. The image A_h is degraded to the low resolution l (A_l, i.e. a SPOT panchromatic image at 20 m) and the images B_{kl} to a lower resolution v (B_{kv}, i.e. a Landsat multispectral image at 60 m), where $v=l(l/h)$,
- the fusion method is applied to these two sets of images, resulting into a set of synthesized images B^*_{kl} at resolution l. The original images B_{kl} (the Landsat original image) serve now as references.
- the second and third properties are tested with the synthetic images B^*_{kl}. The quality observed for the fused products B^*_{kl} is assumed to be close to the quality that would be observed for the fused products B^*_{kh} if a reference at resolution h were present. This point has been largely discussed by Wald *et al.* (1997). A comparison is performed between B_{kl} and B^*_{kl} by the means of visual analysis and analysis of the similarities and discrepancies. For the third property, the emphasis is put on the spectral similarities.

Depending upon the objectives of the assessment and of the available resources, the task of visual analysis will be more or less sophisticated and the computer analysis of the similarities and discrepancies will be more or less extensive. It is also recommended to apply the technique under examination on images of urban areas since these landscapes offer a very large diversity of both spectral signatures and high spatial frequencies.

The protocol followed by Liu is based upon the comparison between the synthetic images B^*_{kh} and the interpolated original images B^{interp}_{kh} (interpolated Landsat images at 10 m). Except in the case of landscapes with no high frequency, the interpolated images B^{interp}_{kh} are different from the images B_{kh} that should be observed if a sensor of resolution h were existent. This has been proven several times. Accordingly, even if the sets of images B^*_h and B^{interp}_h are very similar, it does not prove the high-quality transformation of the multispectral content when increasing the spatial resolution performed by the SFIM method because the reference image B^{interp}_h selected by Liu is not at all the appropriate one.

A further comment is the choice made by Liu of the linear correlation coefficient. This coefficient is not sufficient to summarize a comparison. It is notably insensitive to a linear transformation, except for the limitations due to the accuracy of the computers. It follows that if an image B^*_{kh} offers a bias and / or a much different signal / radiance dynamics than the reference image, it will not appear in the correlation coefficient. Bias and root mean square error (or standard-deviation) are better representing the errors. Furthermore, the correlation coefficient is sensitive to the very high frequencies. This is a surprising choice knowing that interpolated images do not contain such frequencies.

4. Conclusion

The protocol used by Liu to demonstrate the quality of the products fused by the SFIM technique is fully inadequate for its part based on statistics. In theory, this is also true for the visual part but the effects are dampened by the combined effects of the computer display and the visual perception. It results that there is no demonstration of the superiority of the SFIM technique on any other technique based on statistics. We encourage Dr. Liu to adopt the protocol discussed here. It has been debated within a joint working group of the European Association of Remote Sensing Laboratories (EARSeL), the European branch of the International Society for Photogrammetry and Remote Sensing (ISPRS) and of the Société des Electriciens et Electroniciens, the French arm of the IEEE. The work of Wald *et al.* (1997) on this protocol has been distinguished by the American Society of Photogrammetry and Remote Sensing. This protocol is already in operation in several companies and agencies. By using this protocol, the demonstration of the advantages of the SFIM technique can be made and comparisons can be made with already published works (Blanc *et al.* 1998; Mangolini *et al.* 1995; Munechika *et al.* 1993; Ranchin, Wald, 1998, 2000b; Raptis *et al.*, 1998; Vrabel, 2000; Yang *et al.* 2000).

REFERENCES

Blanc Ph., Blu T., Ranchin T., Wald L., Aloisi R., 1998, Using iterated rational filter banks within the ARSIS concept for producing 10 m Landsat multispectral images. *International Journal of Remote Sensing*, 19, 12, 2331-2343.

Liu, J. G., 2000, Smoothing filter-based intensity modulation: a spectral preserve image fusion technique for improving spatial details. *International Journal of Remote Sensing*, 21(18): 3461:3472.

Liu, J. G. and Moore, J. M., 1998, Pixel block intensity modulation: adding spatial detail to TM band 6 thermal imagery. *International Journal of Remote Sensing*, 19(13):2,477-2,491.

Mangolini, M., T. Ranchin, and Wald, L., 1995, Evaluation de la qualité des images multispectrales à haute résolution spatiale dérivées de SPOT. *Bulletin Société Française de Photogrammétrie et Télédétection*, 137, pp. 24-29.

Munechika, C. K., J. S. Warnick, C. Salvaggio, and Schott, J. R., 1993, Resolution enhancement of multispectral image data to improve classification accuracy. *Photogrammetric Engineering & Remote Sensing,* vol. 59, 1, pp. 67-72.

Pradines, D., 1986, Improving SPOT image size and multispectral resolution. *Earth Remote Sensing using the Landsat Thematic Mapper and SPOT Systems, Proceedings SPIE Conference*, 660, pp. 78-102.

Ranchin T., Wald L., 1998. Sensor fusion to improve the spatial resolution of images : the ARSIS method. In : Proceedings, EARSeL Symposium 1997 “ Future Trends in Remote Sensing ”, Lyngby, Denmark, P. Gudmansen Ed., A. A. Balkema Publ., pp. 445-451.

Ranchin, T., and Wald, L., 2000a, Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation. *Photogrammetric Engineering & Remote Sensing*, 66(1), 49-61.

Ranchin, T. and Wald, L., 2000b, Comparison of different algorithms for the improvement of the spatial resolution of images. *In* Proceedings of the third conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images", Sophia Antipolis, France, January 26-28, 2000, Thierry Ranchin and Lucien Wald Editors, published by SEE/URISCA, Nice, France, pp 33-41.

Raptis, V. S., R. A. Vaughan, T. Ranchin, and Wald, L., 1998, Assessment of different data fusion methods for the classification of an urban environment. *In* Proceedings of the second conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images", Sophia Antipolis, France, January 28-30, 1998, T. Ranchin and L. Wald Editors, published by SEE/URISCA, Nice, France, pp. 167-182.

Vrabel J., 2000, Multispectral imagery advanced band sharpening study. *Photogrammetric Engineering & Remote Sensing*, 66, 1, 73-79.

Wald, L., T. Ranchin, and Mangolini, M., 1997, Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images. *Photogrammetric Engineering & Remote Sensing*, 63(6): 691-699.

Wald, L., 2001, *Fundamentals in data fusion and image fusion techniques*. To be published by Les Presses, Ecole des Mines de Paris, Paris, France.

Yang W., Cauneau F., Paris J.-P., Ranchin T., 2000, Influence of landscape changes on the results of the fusion of P and XS images by different methods. *In* Proceedings of the third conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images", Sophia Antipolis, France, January 26-28, 2000, Thierry Ranchin and Lucien Wald Editors, published by SEE/URISCA, Nice, France, pp 47-56.