Mutual Information Restoration of Multispectral Images Using A Generalized Neighborhood Operation

Hesan Z. Rafi1, Hamid Soltanian-Zadeh1, 2 1 Control and Intelligent Processing Center of Excellence
Electrical and Computer Engineering Department
University of Tehran, Tehran 14395-515,Iran
Email:hesanrafi@yahoo.com 2 Medical Image Analysis Lab.
Radiology Department
Henry Ford Health System
Detroit, MI 48202,USA
Email: hamids@rad.hfh.edu

No Institute Given

Abstract. Information theory based techniques for signal and image processing are now considered as a viable alternative to other popular techniques. This paper extends our pervious which was a new multispectral filter based on mutual information maximization which mutually restores multispectral images. For the sake of simplicity we consider only two multispectral images, but the idea can be generalized to more images. Since multispectral images contain analogous information about a scene, as a rule thier mutual information is assumed to be maximal; but noise and other independent artifacts decrease their mutual information. As an extension to our pervious work we have used a generalized neighborhood operation based on an alternative mutual information measure to increase the mutual information between the two neighborhood windows, sliding simultaneously on both images.The main feature of this generalized neighborhood operation is that, it updates all pixels inside the neighborhood window. This filter does not assume any specific relation among the gray level intensities of images, and uses both inter-frame and intra-frame information to suppress noise. Another important modification in this work is addition a new term which filters images individually. This leads to suppress correlated noise and artifacts. Application of the proposed method to simulated images shows the outperformance of this method compared with Perona-Malik method which has received much attention in recent years because of its capability in both noise reduction and edge enhancement.

1 Introduction

Information fusion of multispectral images is a very important issue in remote sensing and medical image analysis. Fusion process submits the multispectral images to some preprocessing steps. Registration (i.e. spatial realignment) of images is the most important phase of preprocessing. As another preprocessing step, images are passed through filters to reduce noise and increase performance of fusion process.

Since filtration is of great importance in image processing, a huge number of filtration methods have been proposed over years. These methods can usually be considered as neighborhood operations on a *single* image [7]. That is to say they are not devised to mutually restore multispectral images. On the other hand,information theory based techniques for signal and image processing are now considered as a viable alternative to other popular techniques.

In this paper, we extend our pervious work which was a new multispectral filter based on Mutual Information (MI) maximization. By definition MI is the amount of information that one variable conveys about the other[5]; therefore, since registered multispectral images are informative of one and the same scene, they should have maximum possible MI, but noise and other independent artifacts decrease MI between them. This filter is a generalized neighborhood operation which increases the MI between two sliding windows of the same coordinate.In contrast to our pervious work neighborhood operation, which updates only the central pixel of an odd sized neighborhood window[1], the new generalized neighborhood operation updates all pixels inside the neighborhood window. This filter does not assume any specific relation among the gray level intensities of images, and uses both inter-frame and intra-frame information to suppress noise. Another important modification with respect to our pervious work is adding a term which tries to restore each image individually. Since we are looking to the images through a small neighborhood window we might have accidently correlated noise in both windows, the filter that we had proposed perviously amplifies these noises and artifacts. This was a main limitation in our pervious filter. This work over comes the limitation by restoring each image individually at a same time.

In section II, we briefly review a relatively new type of MI measure which has been proposed by Xu *et al* [2] and enables us to estimate the MI of two small data sets using a closed mathematical formula directly through data. In section III, we describe the new generalized neighborhood operation which filters multispectral images. Finally, in section IV, we present experimental results and compare the new proposed filter with Perona-Malik filter which is widely used to preprocess multispectral images before multispectral segmentation.[8] this filter has received much attention in recent years because of its capability in both noise reduction and edge enhancement.[7].

2 Alternative Mutual Information (MI) Measures

The MI maximization has been successfully used by Viola [3] to register multispectral images. If multispectral images are not registered, their MI decreases. This registration process finds a transform that maximizes the MI between images. As Viola has pointed out, estimating Shannon's MI by pdfs is an inordinately difficult task. So he estimates Shannon's MI using sample mean method which requires a large amount of data. This estimation method is not suitable to estimate the MI between small data sets like two neighborhood windows where we have a small amount of data and we need to know the influence of each sample on the overall MI.

In this section we briefly review a relatively new alternative MI measure proposed by Xu *et al* [2] which enables us to estimate the MI between two small data sets directly through data samples using a closed mathematical formula.

MI measures the relationship between two variables; in other words, MI is the measure of uncertainty removed from one variable when the other is given. Following Shannon [4],[5] the MI between two RV's X_1 and X_2 is defined as

$$I_s(X_1, X_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2) \log \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1) f_{X_2}(x_2)} \mathrm{d}x_1 \mathrm{d}x_2 \quad (1)$$

This measure could also be regarded as the Kullback-Lieber divergence between the joint pdf $f_{X_1,X_2}(x_1, x_2)$and the factorized marginal pdf's $f_{X_1}(x_1)f_{X_2}(x_2)$.The Kullback-Lieber divergence between two pdfs *f(x)* and *g(x)* is defined as

$$D_{KL}(f,g) = \int_{-\infty}^{\infty} f(x) \log \frac{f(x)}{g(x)} \quad (2)$$

As pointed out by Kapur[6] there is no reason to restrict MI only to this distance measure. Another possible distance measure is based on Cauchy-Schwartz inequality

$$D_{CS}(f,g) = \log \frac{(\int_{-\infty}^{\infty} f(x)^2 \mathrm{d}x)(\int_{-\infty}^{\infty} g(x)^2 \mathrm{d}x)}{(\int_{-\infty}^{\infty} f(x) g(x) \mathrm{d}x)^2} \quad (3)$$

Obviously, $D_{CS}(f, g) \geq 0$ with equality iff f(x)=g(x) almost everywhere. Thus with D_{CS} as a measure of distance, we may define Cauchy-Schwartz Quadratic Mutual Information (CS-QMI) between two variables X_1 and X_2 as

$$I_{CS}(X_1, X_2) = D_{CS}(f_{X_1,X_2}(x_1, x_2), f_{X_1}(x_1) f_{X_2}(x_2)) \quad (4)$$

Therefore, for the given data set $\{a(i) = (a_1(i), a_2(i))^T | 1 \leq i \leq N\}$ of a random variable $X = (x_1, x_2)$ to estimate the CS-QMI of x_1 and x_2 we must estimate the joint and the marginal pdf's of x_1 and x_2. Parzen window method with Gaussian kernel is used to estimate these pdfs.

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{G}(x_1 - a_1(i), \sigma^2) \boldsymbol{G}(x_2 - a_2(i), \sigma^2)$$

$$f_{X_1}(x_1) = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{G}(x_1 - a_1(i), \sigma^2) \tag{5}$$

$$f_{X_2}(x_2) = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{G}(x_2 - a_2(i), \sigma^2)$$

Where $G(x, \sigma^2)$ is a Gaussian kernel. Using the following identity where a and b are considered to be constants

$$\int_{-\infty}^{\infty} \boldsymbol{G}(x - a, \sigma^2)\boldsymbol{G}(x - b, \sigma^2)\mathrm{d}x = \boldsymbol{G}(a - b, 2\sigma^2) \tag{6}$$

we have

$$I_{CS}(X_1, X_2) = \log \frac{V_J\ V_M}{V_C^2} \tag{7}$$

where

$$V_J = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} \boldsymbol{G}(a_1(i) - a_1(j), 2\sigma^2)\boldsymbol{G}(a_2(i) - a_2(j), 2\sigma^2)$$

$$V_M = [\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} \boldsymbol{G}(a_1(i) - a_1(j), 2\sigma^2)]\times [\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} \boldsymbol{G}(a_2(i) - a_2(j), 2\sigma^2)] \tag{8}$$

$$V_C = \tfrac{1}{N}\textstyle\sum_{i=1}^{N}\{[\tfrac{1}{N}\sum_{j=1}^{N} \boldsymbol{G}(a_1(i) - a_1(j), 2\sigma^2)]\times [\tfrac{1}{N}\sum_{j=1}^{N} \boldsymbol{G}(a_2(i) - a_2(j), 2\sigma^2)]\}$$

So we can easily estimate the I_{CS} of two variables using these formulae. For further study about this MI measure and some other generalized information measures and their applications, interested reader is referred to [2].

3 Mutual Information Maximization using Generalized Neighborhood Operations

Neighborhood operations are the central tools for low level image processing. These operations are used to extract certain features from an image.That is why the image resulting from a neighborhood operation is also called a feature image. Proper combination of neighboring pixels can perform quite different image processing tasks such as detection of simple local structures (i.e. edges, corners, lines), motion determination, reconstruction of images taken with indirect imaging techniques (tomography), and restoration [7].

The most important characteristic of a neighborhood operation is the size of the neighborhood window. Although neighborhood operations are usually defined on a single image [7],we can generalize neighborhood operations to N images.

For the sake of simplicity, we consider the case $N = 2$ and select a 3×3 neighborhood window. We can consider these two multispectral images as a single image whose pixels are 2×1 vectors. The vectors contain the gray level of multispectral pixels with the same coordinates.

As shown in Fig.1, each neighborhood window is actually composed of two neighborhood windows sliding simultaneously over both images. Since they ideally contain the same information, their MI should be maximal; but noise causes a decrease in MI. Increasing only the MI like this causes to amplifying the accidental correlated noise, therefore we increase the term which decreases the grey level entropy in the window to restore each image individually at a same time. The same notations as the pervious section are used to denote the gray level values of the pixels inside the windows; therefore the MI between these two windows can be calculated using (7).

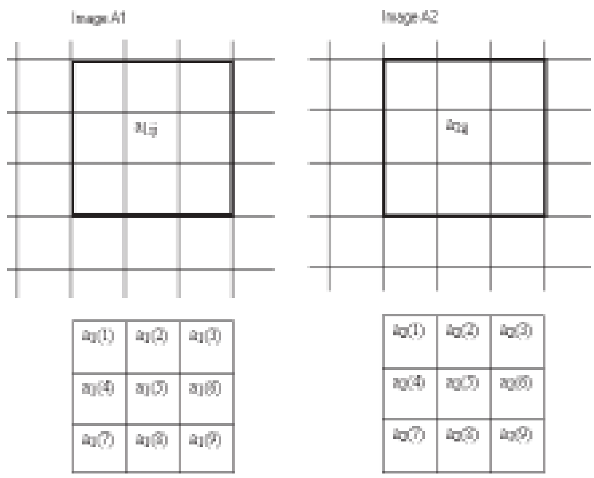

Fig. 1. Two neighborhood windows with the same coordinates slide over multispectral images A_1 and A_2, and the neighborhood operation increases MI between these neighborhood windows

In conventional neighborhood operations, we try to estimate the true value of a pixel from its surrounding pixels.In fact neighborhood operations make use of pixel dependencies in a small region of image (neighborhood window), and since the pixel dependencies are supposed to be isotropic, with regard to symmetry the central pixel of an odd sized neighborhood window receives the result of operation.But as we want to increase the MI of the gray level in a neighborhood window there is no logical reason to restrict modification only to the central pixel. We also

Therefore we intend to modify the pixels gray level in order to increase the MI and decrease the gray level entropy. The gradient descent method can be used to increase MI and decrease the gray level entropy; therefore, the derivative of MI with respect to each pixel is calculated as follows

$$\frac{\partial I_{CS}}{\partial a_k(p)} = \frac{\partial V_J}{\partial a_k(p)}\frac{1}{V_J} + \frac{\partial V_M}{\partial a_k(p)}\frac{1}{V_M} - 2\frac{\partial V_C}{\partial a_k(p)}\frac{1}{V_C} \tag{9}$$

$$k = 1, 2$$

$$1 \leq p \leq 9$$

The new values for the pixels are calculated by adding the initial values and (9) multiplied by a coefficient known as learning coefficient which plays an important role in maximization process. Choosing values greater then 0.5 usually causes the MI oscillate around its initial value. The second term in (9) is a term which decreases the gray level entropy [2].We add this term in order to restore each image independently.

$$a_{k,new}(p) = a_{k,old}(p) + \gamma\frac{\partial I_{CS}}{\partial a_k(p)} + \gamma\frac{1}{V_M}\frac{\partial V_M}{\partial a_k(p)}$$

By scanning the whole image using this neighborhood operation the MI between the feature images increases and, as a result, the two images become mutually restored. Since we have used the gradient descent method to maximize the MI and minimize the gray level entropy between windows, we should iteratively subject the resulting feature images to this filter; in other words this new filter is an iterative filter.

4 Experimental Results

In this section, we compare the experimental results of the new proposed filter with that of Perona-Malik filter which has received much attention in recent years, and is widely used as a preprocessing in many multispectral segmentation methods [7].

Perona-Malik's filter achieves both noise reduction and edge enhancement through use of an anisotropic diffusion equation which in essence acts as an unstable inverse diffusion near edges and as a stable linear-heat-equation -like diffusion in homogeneous regions.

This filter has been implemented using neighborhood operations. Considering the first image A_1 and its corresponding neighborhood window in Fig. 1 the central pixel is updated according to following formula [9]

$$\triangle a_1(5) = \gamma \frac{\sum_{i=1}^{9} (a_1(j) - a_1(i)) e^{\frac{-(a_1(j)-a_1(i))^2}{k^2}}}{\sum_{i=1}^{9} e^{\frac{-(a_1(j)-a_1(i))^2}{k^2}}} \tag{10}$$

Although Perona-Malik have used a 4-neighborhood system and have announced that an 8-neighborhood window does not significantly change results, we implemented their filter with an 8-neighborhood window. In order to compare these two methods, we have generated two simple simulated multispectral images which represent the same scene as shown in Fig. 2-a and Fig. 2-b. To model the noise generated by the imaging system, we have added a white Gaussian noise to images Fig. 2-a and Fig. 2-b resulting in images Fig. 3-a and Fig. 3-b. The SNRs of these degraded images are respectively 3.56 and 3.54. Fig. 4-a and Fig. 4-b show the restored images using Perona-Malik method after 12 iterations. The SNRs of these restored images are respectively 8.65 and 8.72. As seen in each image, the low contrast edges have been severely diminished. Figures 5-a and 5-b show the results of the implementation of the proposed filter after the same number of iterations. The SNRs of Fig. 5-a and Fig. 5-b are respectively 14.13 and 14.48.

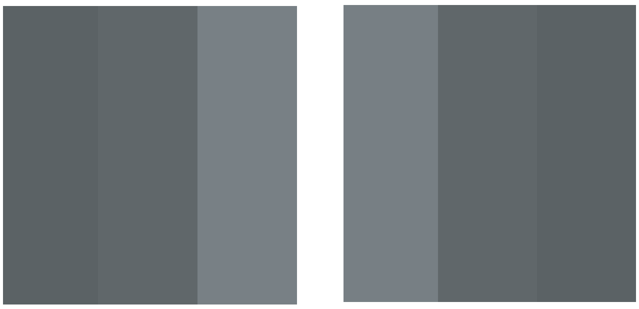

Fig.2-a Fig.2-b

Fig. 2. Two simulated multispectral images representing the same scene.

It can be seen from the images that our method has preserved the edges better than Perona-Malik method. Of course it should be noted that the images are generated in such a way that the edge with low contrast in one image has a high contrast in the other image. Since Perona-Malik method restores images independently it fails to preserve these low contrast edges, but the proposed method may preserve them since it uses both mutispectral images to restore them simultaneously.

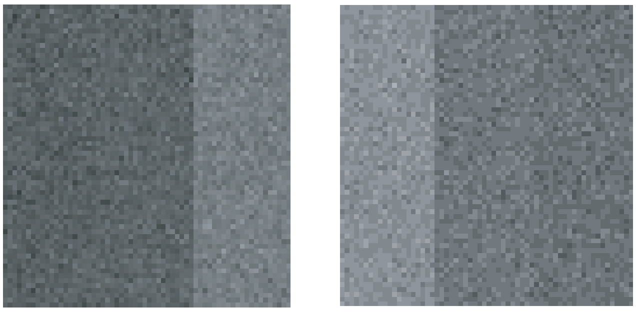

Fig3-a Fig3-b

Fig. 3. Multispectral images degraded by noise.

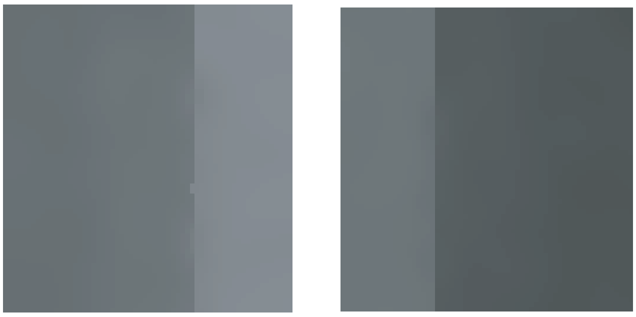

Fig4-a Fig4-b

Fig. 4. Mutispectral degraded images restored by Perona-Malik method.

5 Conclusion

We have proposed a new multispectral filter based on an alternative MI measure using a generalized neighborhood operation. One of the main advantages of this filter is using both inter and intra frame information to suppress noise. The other advantage of this filter is that,there is no necessity to select an odd sized window for this generalized neighborhood operation.This filter also restores each image individually to over come correlated noises. Experimental results show that the SNR of the restored images using the proposed method is higher than that of the Perona-Malik method.

References

1. H. Z. Rafi and H. Soltanianzadeh ,*"Mutual Information Restoration of Multispectral Images"* to be Published in `Proc.IWSSIP`, 2003
2. D. Xu ,*"Energy, Entropy and Information Potential in NeuroComputing"* Ph.D. Dissertation, U. of Florida, 1998.

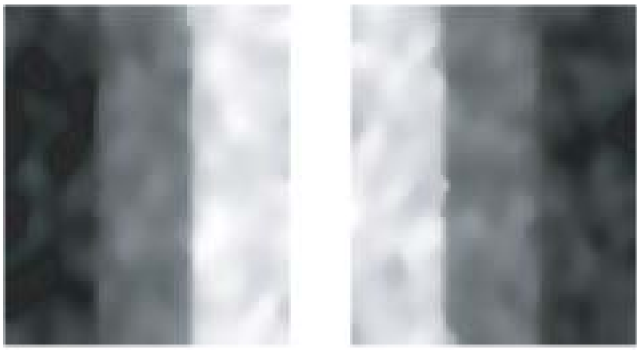

Fig. 5. Multispectral degraded images restored by the proposed method.

3. P. Viola,N. Schraudolph ,T. Sejnowski , *"Empirical entropy manipulation for real-world problems"* Proc. Neural Info. Proc. Sys. (NIPS 8) Conf., 851-857, 1995.
4. C.E. Shannon,*"A mathematical theory of communication"*Ḃell Sys. Tech. J. 27, 1948,pp379-423, 623-653
5. C. Shannon and W. Weaver ,*"The mathematical theory of communication"*, University of Illinois Press, 1949.
6. J.N. Kapur,*"Measures of Information and Their Applications"*J̇ohn Wiley&Sons. 1994
7. B. Jahne,*"Digital Image Processing"* 5^{th} revision. Springer 2002
8. I.N. Bankman *"Handbook of Medical Images Processing and analysis"* Academic Press , 2000
9. P. Perona and J. Malik *"Scale-Space and Edge Detection Using Anisotropic Diffusion"* IEEE Trans. PAMI Vol 12, No 7, July 1990,pp 629-639