Cloud Removal from Satellite Images using Auto Associative Neural Network and Stationary Wavelet Transform

Tapasmini Sahoo
National Institute of Technology
Rourkela, India
tapasmini.sahoo@gmail.com

Suprava Patnaik
S.V.N.I.T
Surat, India
suprava_patnaik@yahoo.com

Abstract

In this paper an image fusion technique is developed to remove clouds from satellite images. The proposed method involves an auto associative neural network based PCAT (principal component transform) and SWT (stationary wavelet transform) to remove clouds recursively which integrates complementary information to form a composite image from multitemporal images. Some evaluation measures are suggested and applied to compare our method with those of covariance based PCAT fusion method and WT-based one. The PSNR and the correlation coefficient value indicate that the performance of the proposed method is better than others. It also enhances the visual effect.

1. Introduction

A significant obstacle of extracting information using satellite imagery is the presence of clouds. Removing these portions of image and then filling in the missing data is an important image editing task. Traditionally the objective is to cut the cloudy portions out from the frame and fill in the gaps with clear patches from similar images taken at a different time. Thus the meticulous work is to identify the cloud and then completing the missing part. Various approaches have been tried to solve this problem with differing results. Wavelet and PCA transformation has showed to be the useful tools to deal with clouds coverage in satellite images.

In this paper the performance of an auto associative neural network based principal components extraction method coupled with wavelet transform is used for image fusion for removal of clouds covering area.

Wang [7] proposed automated cloud detection in a set of two temporal Land sat TM images by simply thresholding high frequency components a extracted by a 2D discrete wavelet transfer of both images. The special frame mentioned in this paper can be called undacimted wavelet transform or stationary wavelet transform (SWT). The SWT is similar to a discrete wavelet transform (DWT) but the difference is that SWT does not perform down sampling of the signal between the levels of hierarchy. For this reason the subimages obtained after the decomposition process have the same resolution as the original images.

Image fusion refers to the technique that integrates complementary information from multiple image sensor data such that the new images are more suitable for the purpose of human visual perception and the computer processing task. In broad sense image fusion is performed at three different processing levels according to the stages at which the fusion takes place, namely pixel level fusion, feature level fusion and decision level fusion. Image fusion at pixel level means that fusion is at lowest processing level. The simplest image fusion on pixel level is to sum and average the original images pixel by pixel. However when this method is applied several undesired effects including reduced contrast of features would appear. If the source level images are RGB color images, the methods of pixel level fusion also include intensity- hue-saturation (IHS) transform and principal component substitution (PCS) and so on. In recent years many researchers recognized that multiscale transforms are very useful for analyzing the information content of images for the purpose of fusion[6], and some sophisticated approaches based on multiscale transforms, such as high pass filtering(HPF) method[5,6], laplacian pyramid[6], morphological pyramid, wavelet transform[1,2] have been proposed. The wavelet transform offers certain advantages over the other above mentioned techniques. It provides directional information and supply spatial orientation in the decomposition. Moreover since wavelet basis functions are chosen orthogonal, the information are each layer of decomposition is unique. Thus it appears to be an efficient information preserving method from

978-0-7695-3267-7/08 $25.00 © 2008 IEEE
DOI 10.1109/ICETET.2008.99

implementation point it incorporates mismatch in contrast of the corresponding components. The proposed algorithm develops a data base of eigenvalues for different parts of the cloudy image and analysis is done.

The paper is organized as follows. In section 2, we briefly introduce PCE technology using auto-associative neural network. The proposed fusion scheme is described in section 3. Section 4 contains the experimental results.

2. PCE using auto-associative neural networks

In this work a new procedure for principal component extraction (PCE) using RLS learning rule has been introduced. Network structure used has G input and output neurons and only K hidden neurons. Network updating can be done sequentially or simultaneously. Let the input vector at time *n* be

$$x(n) = \left[x_1(n), x_2(n), ..., x_G(n)\right]^T \quad \ldots\ldots\ldots\ldots \quad (1)$$

and the weight vector corresponding to the m^{th} neuron be

$$wm(n) = \left[w_{m1}(n), w_{m2}(n), ..., w_{mG}(n)\right]^T \quad \ldots \quad (2)$$

m=1,……,K

Then the output of this neuron can be written as

$$hm(n) = w_m^t(n-1)x(n) \quad \ldots\ldots\ldots\ldots\ldots \quad (3)$$

Assuming that all $1-m$ previous neurons have already been trained and that their weights have converged to the optimal weight vectors $\hat{w}_i, i\varepsilon[1, ..., m-1]$. Then the corresponding extracted principal component are given by

$$y_1(n) = \hat{w}_i^t h(n) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

To extract the m^{th} principal component in the output of m^{th} neuron, the updating model for this neuron should be constructed so that the desired output at iteration *n* is

$$d_m(n) = x(n) - \sum_{i \angle m} \hat{h}_i(m)\hat{w}_i \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

In other words, the neuron must model the network which seeks to generate the original input less all the previously computed $1-m$ outputs. If we define the deflated desired output matrix as

$$D_m(n) = \left[d_m(1), d_m(2), ..., d_m(n)\right]^T \quad \ldots\ldots\ldots \quad (6)$$

then, the optimal weight vector that minimizes the error function

$$J_m(n) = \sum_{k=1}^{m} \left(d_m(k) - h_m(k)w_m(n)\right)^t \left(d_m(k) - h_m(k)w_m(n)\right) \ldots (7)$$

can be obtained in a manner as

$$\hat{w}_m(n) = (h_m^T(n)h_m(n))^{-1} D_m(n)h_m(n) \ldots\ldots\ldots \quad (8)$$

where

$$h_m(n) = \left[h_m(1), h_m(2), ..., h_m(n)\right]^T \quad \ldots\ldots\ldots\ldots \quad (9)$$

Now the RLS algorithm can be applied to compute the optimal weight vector $\hat{w}_m(n)$ at each training sample. The RLS learning rule for the m^{th} node can be summarized as

$$h_m(n) = w_m^t(n-1)r(n) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (10)$$

$$K_m(n) = \frac{P_m(n-1)h_m(n)}{\left[1 + h_m^2(n)P_m(n-1)\right]} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (11)$$

$$e_m(n) = d_m(n) - h_m(n)\hat{w}_m(n-1) \quad \ldots\ldots\ldots\ldots\ldots \quad (12)$$

$$\text{where } d_m(n) = r(n) - \sum_{i<m} \hat{h}_i(n)\hat{w}_i(n-1) \ldots\ldots\ldots \quad (13)$$

$$\hat{w}_m(n) = \hat{w}_m(n-1) + K_m(n)e_m(n) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (14)$$

$$P_m(n) = \left[1 - K_m(n)h_m(n)\right]P_m(n-1) \quad \ldots\ldots\ldots\ldots\ldots \quad (15)$$

where, $P_m(n)$ is the inverse of the covariance of the output of the m^{th} neuron and $K_m(n)$ is the gain for the updating equation of this neuron. Thus the algorithm combines the basic RLS algorithm with the Gram- Schmidt orthogonalization procedure in a manner similar to that of the Generalized Hebbian Algorithm (GHA). Orthogonalization is achieved by subtracting the already determined higher order components from the original input and using the resultant process as a mapping target for the m^{th} neuron. This deflation procedure would implicitly make the effective input to the present neuron equal to the sum of the lower order components associated with eigenvalues $\lambda_m, ..., \lambda_G$. Updating of w_m is then able to extract he significant component λ_m. This component would be orthogonal to the $1-m$ previous higher order component. Thus the number of neurons in the hidden layer can be set equal to K that extracts K most significant components.

Fig.1 describes the process more elaborately. Here we have considered non-overlapping 4 x 4 blocks from an original image and the input vector has been obtained by row /column transformation of the block coefficients. Hidden layer consists of 8 neurons thus the algorithm considers only significant information associated with principal sixteen eigenvectors. For each 4 x 4 block the eight outputs from the hidden

layer neurons are considered for image fusion.

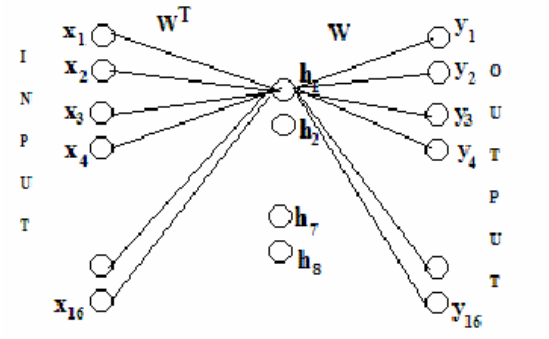

Figure 1. Structure of auto-associative neural network.

In the training phase input vectors are considered from a registered image of some part of Hamirpur area of India. Training phase is for only one pass of the full image. That is if the image is of size 512 x 512 and input vector size is16 (obtained from 4 x 4 sub-block), training is done for 128 x 128 number of samples. After the training phase the weight matrix which corresponds to the eigenvalue of the input image are stored in a data base for analysis of the images taken from the same area.

3. Proposed fusion scheme for removal of cloud

This method merges the detail information of the cloud free image and also preserve the spectral information of the first principal component of the cloud contaminated image. In this approach of removing cloud; fusion is done at pixel level A prerequisite for successful image fusion is that the original images have to be registered correctly so that the corresponding pixels are co-aligned. Figure 2 is a framework illustration of image fusion based on PCA and stationary wavelet transform

step 1: Let the cloud free image is f(x,y) and the first principal component obtained using RLS learning rule of the cloud contaminated image is g(x,y). In the initial step the second image is matched with g(x,y) by the following equation and analyzing the histograms of the two images

$$g(x,y)=\frac{\sigma_g}{\sigma_f}(f(x,y)-\mu_f)+\mu_g \quad \ldots\ldots\ldots (16)$$

where μ_f is the mean of the cloudy image and σ_f is standard deviation of the respective image . similarly μ_g and σ_g is the mean and standard deviation of the first principal component g(x,y).

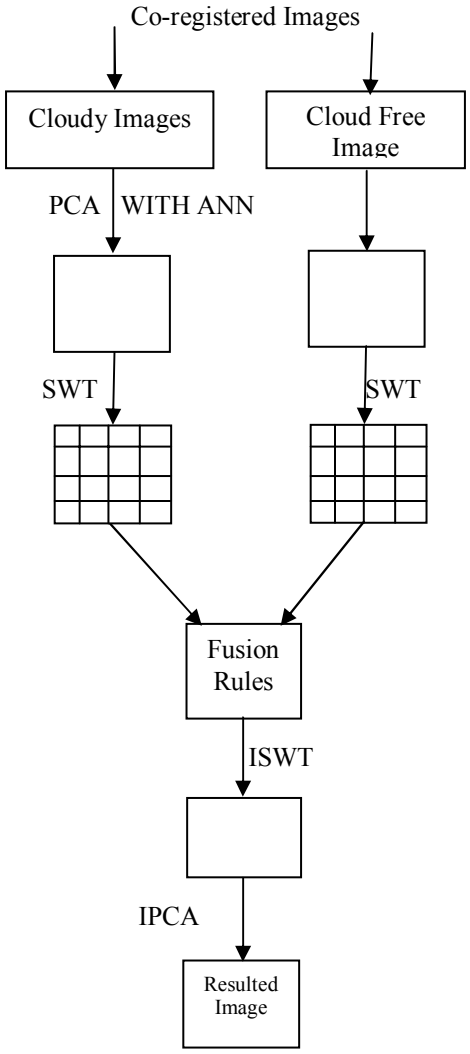

Figure 2. Block diagram of PCA-ANN and SWT based image fusion.

step 2: The two images are then transformed with stationary wavelet, so we obtain low frequency approximate parts $S^k f(2^j;x,y), S^k g(2^j;x,y)$ and high frequency detail parts as $W_k^s f(2^j;x,y), W_k^s g(2^j;x,y)$. J denotes the maximum decomposition level, k=1,2,3,4…. four decomposition parts of the resolution s.

step 3: Let $D_k^s f^j$ and $D_k^s g^j$ are variances of $W_k^s f(2^j; x, y)$ and $W_k^s g(2^j; x, y)$ respectively. We merge the first principal component of the cloudy image with the cloud free image in such away that at the level 2^j, the high frequency detail parts after fusion is

$$W_k^s(2^j; x, y) = \begin{cases} W_k^s f(2^j; x, y) & D_k^s f^j > D_k^s g^j \\ W_k^s g(2^j; x, y) & D_k^s f^j \le D_k^s g^j \end{cases} \quad (17)$$

and the low frequency approximate part is

$$S^k(2^j; x, y) = k1 S^k f(2^j; x, y) + k2 S^k g(2^j; x, y) \quad (18)$$

where k1, k2 are the weighted coefficients and they satisfy the equation k1+k2=1.

Step 4: Eventually the resulted image was obtained by transforming all principal components of the fused image by inverse PCA transform using the same learning rule.

4. Experimental results and performance comparison

This approach takes into account at least two satellite images of the same area at different times. As a preprocessing condition, those images have to be co-registered using image to image registration process. This process ensures that the same pixel in both images refers to the same area. Then the two images are merged as per the fusion process explained in fig 2. The two, cloudy and cloud free images are of size 512 × 512 ; illustrated in fig.3 and fig 4 respectively. The fused image using the proposed method is shown in fig 5, and the fused image using traditional covariance based PCA transform method is fig 6; fig 7 is the fused image using pixel based WT fusion method. Some statistical parameters like PSNR, deviation constant and correlation coefficient are calculated as a factor of performance evaluation. Definition of those parameters are stated as follows

1) PSNR : Quality assessment of fused images are traditionally carried out by visual analysis. So in this work we have taken peak to peak signal-to-noise ratio (PSNR). As pr this if R is the standard reference image and F is the fused image, the root mean square error is defined as follows

$$RMSE = \sqrt{\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left[R(i,j) - F(i,j)\right]^2}{M \times N}} \quad \ldots\ldots\ldots \quad (19)$$

Figure 3. Image with cloud.

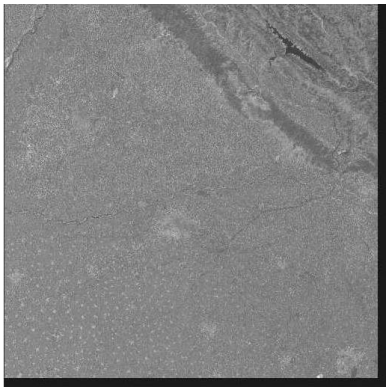

Figure 4. Cloud free image.

Figure 5. Fused image with the proposed algorithm.

and corresponding PSNR is

$$PSNR = 10 \times \ln\left(\frac{p_{\max} \times p_{\max}}{RMSE^2}\right) \ldots\ldots \quad (20)$$

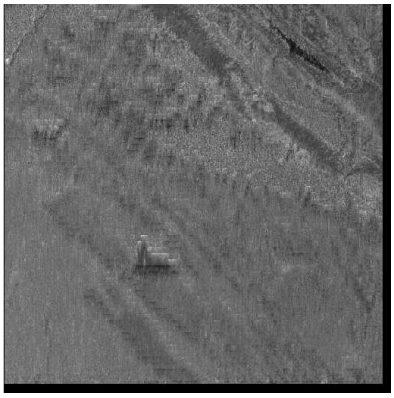

Figure 6. Fused image with pixel based WT.

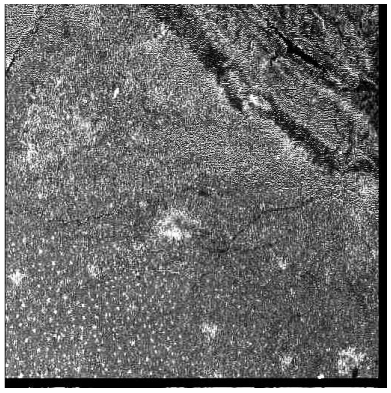

Figure 7. Fused image with covariance based PCA.

where $p_{\max}$ is the maximum gray value of the pixels in the fused image. Thus bigger the value of PSNR, better the fusion process.

2) Deviation Index: It is defined as follows

$$D_{index} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{\left|I(i,j)-I'(i,j)\right|}{I(i,j)} \quad \text{.....(21)}$$

where I denotes the intensity value of the original cloud free image and I' denotes the intensity value of the fused image in the position (i,j).The lesser the value of deviation index, better the fusion process.

3)Correlation Coefficients: It reflects the correlation extent between the original and fused image. The two images' correlation coefficients is defined as

$$C(f,g) = \frac{\sum_{i,j}\left|(f_{i,j}-\mu_f)\times(g_{i,j}-\mu_g)\right|}{\sqrt{\sum_{i,j}(f_{i,j}-\mu_f)^2\times\sum_{i,j}(g_{i,j}-\mu_g)^2}} \quad (22)$$

where $f_{i,j}$, $g_{i,j}$ denotes the gray value in the position (i,j) of the two images and μ_f, μ_g is the mean of the two images respectively.

Table 1. Performance statistical parameter

Method	PSNR	Correlation Coefficient	Deviation Index
PCA with covarince	89.213	0.871	0.069
Wavelet (SWT)	92.5622	0.923	0.063
PCA with ANN and SWT	94.376	0.946	0.058

5. Conclusion

The proposed algorithm preserves the spectral information in the fused image. This approach also merges the spatial details of the two images. So the performance of this fusion method is improved greatly as compared to other methods. The new algorithm not only enhances the fused image's ability to express the spatial details but also can preserve spectral information of the cloudy image.

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts if possible. True-Type 1 fonts are preferred.

6. References

[1] H.H. Wang, "A new multiwavelet based approach to image fusion", *Journal* Mathematical Imaging and Vision, v1.21, pp.177-192, Sep 2004.

[2] G. Pajares, J.M. l de la Cruz, " A wavelet-based image fusion tutorial", Pattern Recognition, vol.37, pp.1855-1872, 2004.

[3] S. Bannour and M.R. Azimi Sadjadi, " Principal Component Extraction using Recursive Least Squares Learning", IEEE Trans. on Neural Network, vol.6, issue 2, pp. 457-469, March 1995.

[4] P.S. Chavez, S.C. Slides and J.A. Aderson, "Comparison of different methods to merge multiresolution and multispectral data", Photogrammetric Engineering and Remote Sensing vol.57, no.3, pp.295-303, Mar. 1991.

[5] A. Toet, "Hierarchical image fusion", Machine vision and application, 3(1) pp.1-11, 1990.

[6] Burt P J, Adelson E H, "The laplacian pyramid as a compact image fusion code", IEEE Trans. on communication, 31(4), pp. 532-540, 1983.

[7] Bin Wang, Atsuo Ono, Kanako Maramatsu, and Noburo Fujiwara, "Automated detection and removal

of clouds and their shadows from landsat TM images. IEICE Trans. Inf. and Syst., e82-d (2), February 1999.

[8] H. Wang, Z. Jing, and J. Li, "Image fusion using non separable wavelet frame". Chinese optics letters, 1(9), September 2003.

[9] S. Wisetphanichkij, K. Dejhan, F. Cheevasuvit, S. Mitatha, and C. Netbut , " Multi-temporal cloud removing based on image fusion with additive wavelet decomposition", Faculty of Engineering and Research Center for Communication and Information Technology.