# A Survey of Fusion Techniques for Multi-spectral Images

Tiranee Achalakul
Department of Computer Engineering, Faculty of Engineering,
King Mongkut's University of Technology Thonburi,
91 Pracha-Uthit Road, Bangkok 10140, Thailand
Tel. +66-2-470-9380, Fax. +66-2-872-5050

**Abstract:** This paper discusses various algorithms to the fusion of multi-spectral image. These fusion techniques have a wide variety of applications that range from hospital pathology to battlefield management. Different algorithms in each fusion level, namely data, feature, and decision are compared. The PCT-Based algorithm, which has the characteristic of data compression, is described. The algorithm is experimented on a foliated aerial scene and the fusion result is presented.

## 1. Image Fusion Theory and Applications

Image Fusion is the process of combining information from a variety of sensors to produce a unified result. The fusion process ideally generates a single color-composite image that represents all the useful information from a set of images of different sensors/wavelengths (multi-spectral image), thus removing the problems inherent in frame-by-frame evaluation. The main goal is to improve data interpretation and recognition by taking advantage of the complementary characteristics of different sensors. The concept of image fusion is shown in Figure 1. [11, 12].

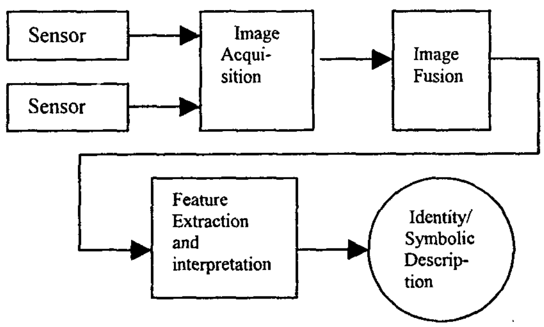


Figure 1. Concept of Multi-sensor Fusion

Objects of different materials normally have unique spectral signatures. In other words, different objects will absorb and reflect differently to light. Using this assumption, we can define a spectral signature for each type of objects in the scene by experimenting with light from different wavelengths. Fusion of a multi-spectral image, thus, can be used to enhance objects of interest in any particular scene. The resulting image can then be used for feature extraction, statistical calculation, and object identifying. This technology has a wide variety of applications that include land-mine detection, monitoring of agricultural resources and the weather, military camouflage detection, computational staining in pathology, and surveillance.

## 2. Fusion Methods for Multi-spectral Images

A multi-spectral image may be fused using various techniques: data level fusion, feature level fusion, and decision level fusion [11].

**Data Level Fusion.** Fusion at this level, a raw image acquired from multiple sensors is fused together on a pixel-by-pixel basis. The algorithms concern fusion in either spatial or frequency domain. In spatial domain, techniques usually involved image arithmetic (addition, subtraction, multiplication, and division) on the pixel intensity from two or more bands. Band differences or band ratios [25] are the most useful of these approaches and are often used to enhance spectral reflectance differences for rocks, soils, and vegetation. Unfortunately, it is unclear how to define effective fusion arithmetic for a large number of bands. Empirical selection of arithmetic rules may introduce losses in pixel contrast, and important spectral information is thus lost. An alternative fusion technique that uses maximum contrast selection [22] may involve a contrast measurement calculation for each pixel at each scale and orientation in all spectral bands.

Fusion in frequency domain, an original image is transformed to frequency or spatial-frequency domain, where contrast and object characteristics are readily available. The method of choice is application-dependent; however, the algorithms typically follow the general schema described in Figure 2.

Source images are usually transformed using Fourier [10, 20], Laplacian [4, 5], or Wavelet transforms [19, 23]. The representations in the transform domain are then combined using algebraic rules to form a single fused data set. This result is then inversely transformed to obtain a final visible image. The Fourier transform, although efficient, does not correlate high frequency components with the original spatial information and hence cannot locate the position of an interesting attribute within an image. This problem has been resolved using pyramid-based methods such as the Laplacian and Wavelet transformations. The Wavelet transform has several advantages over the Laplacian: it provides a more compact representation, separates spatial orientation in different bands, and decorrelates interesting attributes in the original image.

The algebraic rules typically weight images in the source because of the relative importance of specific spectra [22]. For example, in concealed weapons detection, images in the millimeter wave spectrum extenuate metallic objects and are weighted higher in order to show them on the background of visible images representing people [29]. The algebraic rules may be based on pixel intensity [21] or on some measure of contrast [28]. This latter concept allows the selection of the spectral band that should dominate in the fused image. One interesting method is the maximum selection rule introduced by Burt [4, 5] for combining the coefficients of Wavelet and Laplacian transforms.

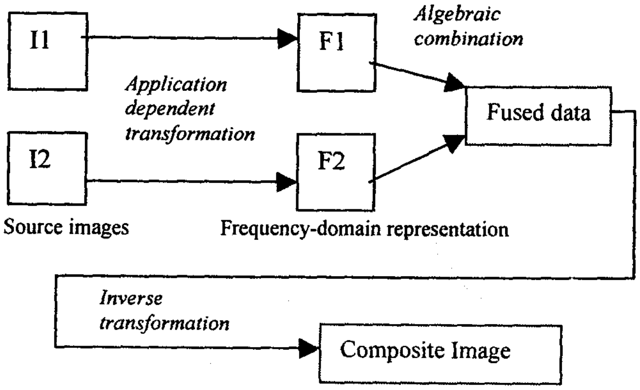


Figure 2. Overview of the Fusion Process

**Feature Level Fusion.** At the feature level, raw data will be transformed in the output into a representation such as image segments or signal amplitude or as shape, length, or orientation of objects in an image. The typical algorithms used in this level are parametric templates [8], hierarchical clusters [7], neural networks [14, 15, 16], and knowledge-based approaches [3]. Parametric templates are often used because of their simplicity. The effectiveness of these methods depends upon the distribution in a feature space: if the distribution is low, the overlap will introduce ambiguity that may not be resolvable. To enhance the result with little added complexity, the hierarchical cluster method can be used. In this method, a cluster is an abstract description of a set of objects in the image that may be divided into sub-clusters by virtue of application-dependent parameters that discriminate objects.

An alternative method is to use Artificial Neural Networks (ANN), which performs a nonlinear transformation between an input vector and an output feature. This method can produce the required output more efficiently than most other approaches [12]. In some medical applications, multiple computed tomography images taken from many different planar views, angles, distances, and spectrums are used as inputs to this algorithm. The output is a computed tomography image that has a clearer view and enhanced quality. Unfortunately, a considerable level of training is required to achieve the transformation function. A training set is made up of ($X_i$, $Y_i$) data pairs where $Y_i$ is a set of output associated with the input vector $X_i$. The association functions between the input and output vectors are generated during the training session. After training is performed, the ANN can construct a novel CT image for any given inputs within seconds. To start the training procedure, the input images are fed through the network, where random interconnection weights are generated. The interconnection weights are adjusted throughout several iterations. The association functions are determined when the root-mean-square error is less than a certain limit. These association functions will later be used to construct output images from any given set of inputs. The root-mean-square error, C, is calculated as follows:

$$C = \{ [1/(N*J)] * [\Sigma\Sigma(D_{n,j} - Y_{n,j})^2] \}^{1/2}$$

where J is the number of nodes, N is the number of patterns in the training set, $D_{n,j}$ is the expected jth output node value for the nth training set pattern, and $Y_{n,j}$ is the actual output value. The interconnection weight values are adjusted as follows:

$$w_{i,j\,new} = w_{i,j\,old} + \eta\sigma_i x_i + \alpha(w_{i,j\,old} - w_{i,j\,previous})$$

where $w_{i,j\,old}$ is the present weight value, and $w_{i,j\,previous}$ is the weight value before being adjusted to $w_{i,j\,old}$; $\sigma_i$ is the error difference in the ith node multiplied by the derivative sigmoid activation function:

$$\sigma_i = f'(y_i)(d_i - y_i)$$

The magnitude factor of each adjustment is $\eta$, while $\alpha$ provides an impetus to the weight adjustment [14, 15, 16]. Several thousand training iterations may be required, and thus a substantive computation is required in the training process.

Knowledge-based approaches are alternatives that emulate the cognitive processes used by humans. These approaches emphasize the use of a set of production rules, frame representations, and computational logic. Unfortunately, considerable training is also required for these techniques [12].

**Decision Level Fusion.** At the decision level, sensor data is processed individually and an identity declaration is performed by applying voting techniques [13], scoring models [12], or other *ad hoc* methods [12]. Voting techniques provide discrimination by a simple majority determination, where the most likely object is detected; scoring models form a weighted sum and determine the maximum weighted score.

The preference for one of these alternative approaches is highly application-dependent. According to a survey produced by Hall and Llinas [12], in the sample space of 30 fusion systems there are over 75 algorithms used; no single algorithm can satisfy all of the needs [2]. However, most the image fusion schemes described have a common problem: they are most effective when used on a small number of images taken from different sensors: for example, cameras with pre-selected IR, UV, or Visible filters. It is less clear how to adapt the algorithms to support a large number of input images.

In the following sections, the paper presents an example application for multi-spectral image fusion. The data used is a 210 spectral-band aerial scene in remote sensing. Due to a large number of bands used, we chose to explore fusion algorithm based on Principal Component Transform (PCT) [1], which has the characteristic of compressing data. PCT has been employed in a variety of remote sensing applications including hyper-spectral data compression, information extraction and fusion, [18] and change detection [6, 9, 24].

## 3. PCT-Based Fusion Scheme

The essence of the PCT idea is to summarize and de-correlate the images by removing redundancy and packing the residual information into a small set of images, termed *principal components*. The components are rank ordered by the magnitude of their variances (eigenvalues); therefore, most of the spectral contrast is pushed forward to the first few components, with an increase in the signal-to-noise ratio of these components.

In our investigations, the important spectral information may contain a large number of images, typically between forty and two hundred. Although the data set possesses different information in each band, there is high inter-band correlation due to the common features in a scene. Since data processing involves a significant degree of redundancy, we prefer the alternative fusion scheme (Figure 3) of the Spectral-screening Principal Component Transform (s-PCT) [1]. The conventional PCT algorithm is an approach that may utilize either a correlation matrix or a covariance matrix to de-correlate the source images and thus remove this redundancy [26, 27]. The correlation matrix tends to prevent features with large numerical values from dominating the resulting bands, and although this produces unbiased eigenvalues, it often distributes variation over a larger number of the resulting components than the covariance matrix. Our goal is to pack as much information as possible into the first few principal components, thus we choose to work with the covariance matrix. Our method performs spectral angle classification [17] on the original image sets. The unique spectral set is then used to form the associated covariance matrix, which characterizes variations in image contrast. The covariance matrix is then used to form, through a linear transformation, a collection of principal components that effectively summarize the most significant variations across all spectra.

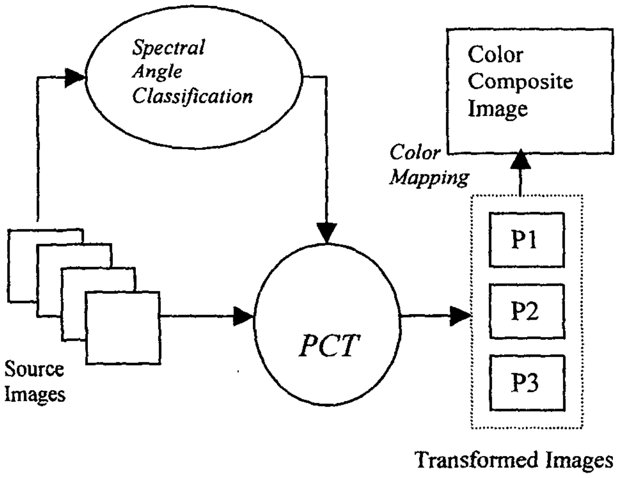


Figure 3. PCT-Based Fusion Algorithm

## 4. Fusion Results

In our experiment we used an aerial scene containing a significant mix of forestry, fields, roads, and a mechanized vehicle. Figure 4 shows a representative sample of frames picked from the 210 spectral bands. The algorithm in section 3 is used to summarize the information from all spectral bands into a single, color-composite image; presenting as much information as possible to human observer.

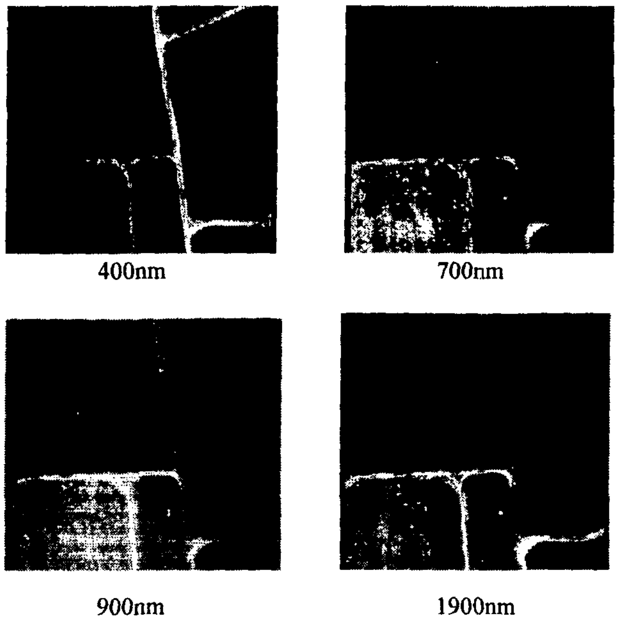


Figure 4: Sample Frames from Original Multi-spectral Image

Figure 5, when viewed on a high-quality monitor, shows improved contrast levels. The forested areas show improved detail and the mechanized vehicle in the lower left corner is significantly enhanced against its background. Post-processing steps can subsequently be applied to detect edges in the image and use structural information to detect and classify the vehicles.

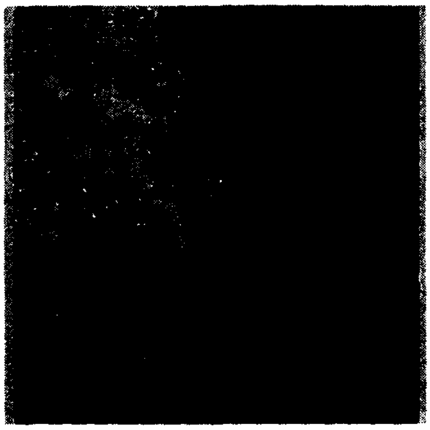

Figure 5: Resulting Image

## 5. Conclusion

This paper has discussed and compared various techniques used for multi-spectral image fusion. Different algorithms are described with the emphasis on the PCT algorithm. PCT-Based algorithms are proved to be effective for data with a large number of spectral bands, such as remote sensing data. An Experiment with an aerial picture of a foliated scene is also presented and the result displays the improvement in image quality.

## References


[1] T. Achalakul and S. Taylor, "A Concurrent Spectral-Screening PCT Algorithm for Remote Sensing Applications", *Journal of Information Fusion*, Vol. 1:2, (2000), pp. 89-97.

[2] D.A. Agarwal, "Totem : A reliable ordered delivery protocol for interconnected local-area networks", Ph.D. Dissertation. Dept. of Electrical and Computer Engineering, University of California, Santa Barbara, 1994.

[3] R.A. Benfer, E. E. Brent, L. Furbee, *Expert System Series: Quantitative Applications in the Social Sciences*, Sage Publication, London, England, 1991.

[4] P.J. Burt, *The pyramid as structure for efficient computation, in Multiresolution Image Processing and Analysis*, Springer-Verlag, New York/Berlin, 1984, pp. 6-35.

[5] P.J. Burt and R. J. Koczynski, "Enhanced image capture through fusion", *Proceedings of the Fourth International Conference on Computer Vision*, Berlin, Germany, pp. 172-182, May 1993.

[6] G.F. Byrne, P. F. Crapper, and K. K. Mayo, Monitoring Land-Cover Change by Principal Component Analysis of Multitemporal Landsat Data, Remote Sensing of Environment, 10, (1980), 175-184.

[7] B. Everett, *Cluster Analysis*, Academic Press, New York, NY, 1980.

[8] K. Fukanaga, *Introduction to Statistical Pattern Recognition*, Academic Press, New York, NY, 1990.

[9] T. Fung and E. LeDrew, Application of Principal Components Analysis to Change Detection, Photogrammetric Engineering and Remote Sensing, 53, (1987), 1649-1658.

[10] R.C. Gonzalez, R. E. Woods, *Digital Image Processing*, Addison-Wesley Publishing Company, Inc., New York, NY, 1993.

[11] D.L. Hall, An Introduction to Multisensor Data Fusion, *Proceedings of the IEEE*, Vol. 85, (1997), pp. 6-23.

[12] D.L. Hall, *Mathematical Techniques in Multisensor Data Fusion*, Artech House, Boston, MA, 1992.

[13] D.L. Hall, A. Nauda, An Interactive Approach for Selecting IR&D Projects, *IEEE transaction on Engineering Management*, Vol. 37:2, (1990), pp. 126-133.

[14] (a) J.P. Kerr, E. B. Bartlett, A Statistically Tailored Neural Network Approach to Tomographic Image reconstruction, *Medical Physics,* Vol. 22:5, (1995), pp. 601-610.

15] (b) J.P. Kerr, E. B. Bartlett, Neural Network Reconstruction of Single-Photon Emission Computed Tomography Images, *Journal of Digital Imaging*, Vol. 8:3, (1995), pp. 116-126.

16] (c) J.P. Kerr, E. B. Bartlett, Medical Image Processing Utilizing Neural Networks Trained on a Massively Parallel Computer, *Computers in Biology and Medicine*, Vol. 25: 4, (1995), pp. 393-403.

[17] F.A. Kruse, A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and F. H. Goetz, The spectral Image Processing System (SIPS) – Interactive Visualization and Analysis of Imaging Spectrometer Data, Remote Sensing Environment, 44, (1993), 145-163.

[18] J. Lee and K. Hoppel, Principal Components Transformation of Multifrequency Polarimetic SAR Imagery, IEEE Transactions on Geoscience and Remote Sensing, 30, (1992), 686-696.

[19] H. Li, B. S. Manjunath, S. K. Mitra, Multisensor Image Fusion Using the Wavelet Transform, *Graphical Models and Image Processing*, Vol. 57:3, (1995), pp. 235-245.

[20] I.M. Martin, D. C. Marinescu, Concurrent Computation and Data Visualization for Spherical-Virus Structure Determination, *IEEE Computational Science & Engineering*, (1998), pp. 40-52.

[21] M. Paval, J. Larimer, and A. Ahumanda, "Sensor Fusion for Synthetic Vision", *AAIA conference: Computing in Aerospace*, Baltimore, MD, October, 1991.

[22] T. Peli, K. Ellis, R. Stahl, "Multi-Spectral Image Fusion for Visual Display", *SPIE vol 3719 Sensor Fusion: Architectures, Algorithms, and Applications III*, Orlando, FL, pp. 359-368, April 1999.

[23] L. Prasad and S. S. Iyengar, *Wavelet Analysis with Applications to Image Processing*, CRC Press, Boston, MA, 1997.

[24] J.A. Richards, Thematic Mapping from Multitemporal Image Data Using the Principal Components Transformation, Remote Sensing of Environment, 16, (1984), 35-46.

[25] J.A. Richards, and X. Jai, *Remote Sensing Digital Image Analysis: An Introduction,* Springer, New York, NY, 1998.

[26] A. Singh and A. Harrison, Standardized Principal Components, *International Journal of Remote Sensing*, Vol. 6, (1985), pp. 883-896.

[27] A. Singh and L. Eklundh, A Comparative Analysis of Standardised and Unstandardised Principal Components Analysis in Remote Sensing, *International Journal of Remote Sensing*, Vol. 14, (1993), pp. 1359-1370.

[28] A. Toet, Multiscale Contrast Enhancement with Application to Image Fusion, *Optical Engineering*, Vol. 31, (1992), pp. 1026-1031.

[29] Z. Zhang and R. S. Blum, "Region-based Image Fusion Scheme for Concealed Weapon Detection", *Proceedings of the 30$^{th}$ CISS Conference*, Baltimore, MD, March 1997.