Hierarchical PCA Techniques for Fusing Spatial and Spectral Observations With Application to MISR and Monitoring Dust Storms

Abhishek Agarwal, Hesham Mohamed El-Askary, Tarek El-Ghazawi, Menas Kafatos, and Jacqueline Le-Moigne

***Abstract*—In this letter, we propose hierarchical principal component analysis (HPCA) techniques for fusing spatial and spectral data, and compare them to direct principal component analysis (DPCA) over Multiangle Imaging SpectroRadiometer (MISR) data. It is shown that the proposed methods are significantly faster than DPCA. In case of DPCA, we merge the 20 different images resulting from the four spectral bands over the nadir and the four forward angles. In the hierarchical case, we first merge the information from the four spectral camera bands; then, we integrate the spatial information from the five cameras in the second step (or vice versa) by applying principal component analysis (PCA) twice. The classification results show that fused data using HPCA compare favorably to DPCA or to classification using the original data. This is because applying PCA to one particular data domain (e.g., spectral data followed by spatial data or vice versa) tends to better remove redundancies and enhance features within that domain. In addition, classification through hierarchical data fusion results in computational savings over the other methods.**

***Index Terms*—Data fusion, dust storms, Multiangle Imaging SpectroRadiometer (MISR), principal component analysis (PCA).**

I. Introduction

DATA fusion usually takes place at three different levels: 1) pixel; 2) feature; and 3) decision [1]. A great variety of fusion methods have been proposed in the literature, and each aims at a different application [2]–[4]. Principal component analysis (PCA) is one of the most well known techniques used in data dimension reduction, yet it has been introduced as a tool for data fusion as well. The principal component (PC) computation involves calculating the covariance matrix of the information in spectral and spatial domains, finding its eigenvalues and eigenvectors, and forming the PCs. By common practice, only the first few components are likely to contain the needed information. PCA condenses all the information of an N-band original data set into a smaller number of new bands than N (or PCs) in a way that maximizes the covariance and reduces redundancy to achieve lower dimensionality. In [5], we have provided a preliminary assessment of using PCA as a data fusion tool, rather than a dimension reduction method, for detecting dust events clearly. Dust storms create major safety and health hazards due to the significant impact on visibility, traffic, and respiratory problems that they create. Hence, timely warnings of dust storms are needed and must be initiated in populated regions for health and traffic control concerns [6]. The Multiangle Imaging SpectroRadiometer (MISR) instrument studies the Earth's climate through the acquisition of global multiangle imagery, which enhances our understanding of the Earth's environment in the third dimension and the fine discrimination between different materials [7]. MISR employs nine discrete cameras pointing at fixed angles, with one viewing the nadir (vertically downward) direction, four viewing the forward, and four more viewing the aftward directions. MISR images are obtained in four spectral bands: 1) blue; 2) green; 3) red; and 4) near infrared. MISR data have been used for dust storm detection in association with other products covering different regions of the electromagnetic spectrum [8]. However, more work is still needed to benefit from the multiangle views over the four wavelengths. In this letter, we attempt to automate this process by means of data fusion to incorporate information from both domains. Spatial and spectral data fusion techniques allow formalizing the integration of these information, as well as monitoring the quality of information in the course of the fusion process [9], [10]. We are using the pixel information that was revealed by the dust event for the fusion process. Thus, we see data fusion as a tool for combining two or more different images to form a new image of greater quality or suitability for the underlying application, which is dust storms in our case. Here, we focus on the ability of PCA to integrate and concentrate information, and to fuse the MISR data, with classification applications in mind. While the primary objective of this technique is speed, the results will also show that this technique leads to better classification results for the monitoring of dust storms using MISR observations. In this letter, we introduce and study PCA techniques that fuse all information and provide spectral data fusion, spatial data fusion, hierarchical spatial-spectral data fusion, and hierarchical spectral–spatial data fusion. In all of these cases, we study the accuracy of the classification of the dust storm as well as the associated computational complexity. We identify and investigate six alternative techniques to perform such PC-based data fusion. In the first baseline technique, we perform no PCA and directly perform classification over the 20 bands that are obtained from the five MISR cameras, where each

Manuscript received November 2, 2006; revised May 14, 2007.

A. Agarwal and T. El-Ghazawi are with the Department of Electrical and Computer Engineering, George Washington University, Washington, DC 20052 USA (e-mail: agarwala@gwu.edu; tarek@gwu.edu).

H. M. El-Askary is with the Department of Environmental Sciences, Alexandria University, Alexandria, Moharam Bek 21522, Egypt and also with the National Authority for Remote Sensing and Space Science (NARSS), Cairo, Egypt (e-mail: helaskar@gmu.edu).

M. Kafatos is with the Center for Earth Observing and Space Research, George Mason University, Fairfax, VA 22030 USA (e-mail: mkafatos@gmu.edu).

J. Le-Moigne is with NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA (e-mail: lemoigne@cesdis.gsfc.nasa.gov).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2007.904467

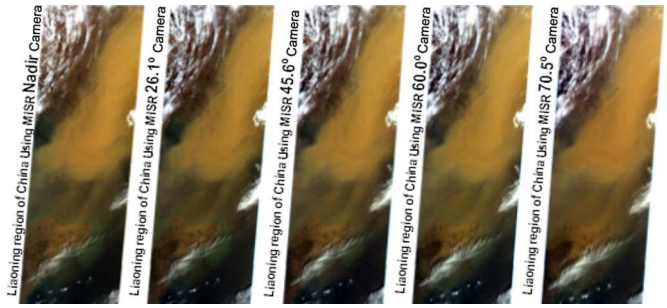

Fig. 1. Different angle views of a large dust plume on April 8, 2002, over the Liaoning region of China and parts of northern and western Korea.

camera produces information at the four bands. In the second technique, we perform DPCA on these 20 bands and use some of the PCs to perform the classification. In the third technique, we perform PCA on each band for the five different angular views (spatial fusion) and then perform the classification on some of the PCs. In the fourth technique, we perform PCA on each camera for the four different bands and then perform the classification on some of the resulting PCs. The fifth and sixth techniques, which are referred to as hierarchical principal component analysis (HPCA) here, are accomplished by performing the PCA on the first PCs that were obtained from the third and fourth methods. To compare the effectiveness of these techniques, the classification results that were obtained by each of them is compared against a reference classification (ground truth). To produce this ground truth, we used the help of an Earth scientist with experience in dust storms to highlight the regions with higher dust concentrations within the dust cloud itself, along with *a priori* knowledge about the dust conditions that were reveled from a news article that was reported in http://visibleearth.nasa.gov/view_rec.php?id=2697.[1]

We have selected a case study representing a dust event over the Liaoning region of China (latitude: 36.3° N–41.3° N, longitude: 121.0° E–123.5° E). MISR multiangle images clearly show a dust storm; however, some cloudy features that can be confused with dust were also observed as in Fig. 1.

Dust storm detection and tracking could be difficult, as dust storms share some similar attributes to those of clouds. Dust storms can vary in their shape, particle size, and distribution; hence, they normally show a varying behavior.

II. Proposed PCA Fusion Technique

When proposing HPCA for data fusion, we identify the other existing different possibilities that were mentioned here for comparison and making the discussion complete. To assess the effectiveness of the proposed methods, we evaluate the quality of the results by classifying the data that were produced from the respective fusion methodology. These data fusion possibilities are stated and discussed here.

A. Existing Techniques Direct Classification

In the direct classification and DPCA methods, as shown in Figs. 2 and 3, we use all the bands that were rendered by the different angle and wavelength images as a multidimensional

[1]Information was last verified on January 3, 2007, at the given website.

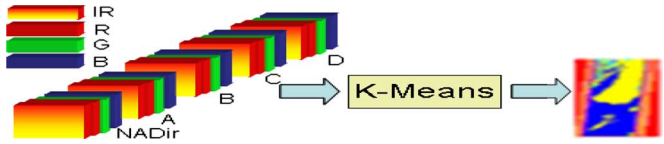

Fig. 2. Direct classification.

Fig. 3. DPCA.

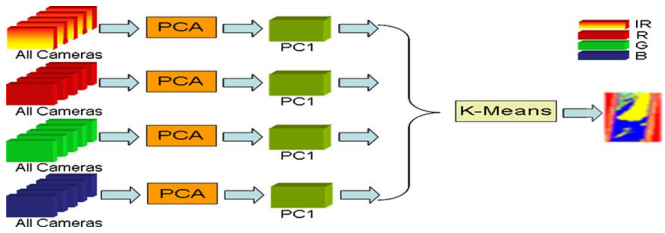

Fig. 4. Spatial PCA.

product with 20 bands (5 angular × 4 spectral). Direct classification using k-means clustering is performed to quantitatively show the prominent sand storm cloud. Since all raw data are used as is, the classification output from this method can be taken as a reference point to gauge the performance of the other techniques. In this case, since the k-means algorithm is performed on all 20 bands, the execution takes a significant amount of time, as compared to performing k-means on the fused data. However, when performing DPCA on the 20 original bands, the first component of the 20 obtained PCs is found to contain about 93.9% of the total data variation (information content), as revealed from the eigenvalues. The classification results were produced using the first component of the reduced data set. The results match the classification results that were obtained from the direct classification method very well. However, the k-means algorithm runs much faster, compared to direct classification. This is because DPCA classification is performed on one band instead of 20 bands as in the direct classification case.These two methods are explained for completeness and to serve as reference methods to compare to the proposed HPCA methods.

B. Spatial and Spectral PCA

In the spatial PCA and spectral methods, as shown in Figs. 4 and 5, we fuse information from the different angles (spatial fusion) for a given camera wavelength and from different wavelengths (spectral fusion) for one particular camera angle, respectively. In spatial PCA, the first eigenvalue of each wavelength contains approximately 94.2% of the total data variance. However, in spectral PCA, the first eigenvalue of each frequency contains approximately 98% of the total data variance. The first components of each method are then combined together, and k-means classification is performed. Classification results that are comparable to those of the direct classification and DPCA methods are observed for both spatial and spectral PCA. Yet, classification in these cases is again much

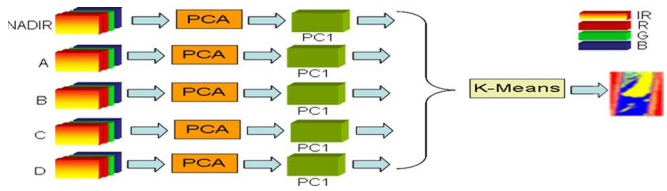

Fig. 5. Spectral PCA.

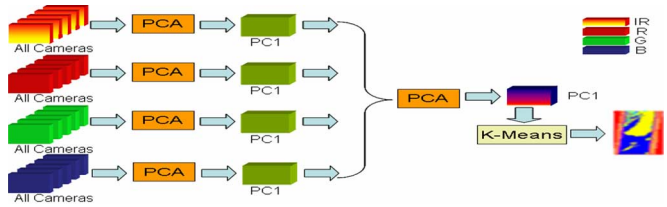

Fig. 6. Spatial–spectral PCA.

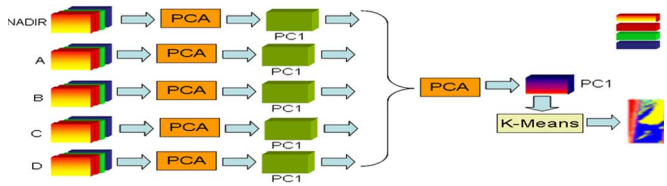

Fig. 7. Spectral–spatial PCA.

faster than the direct classification and slower than DPCA. This is because classification for spatial PCA uses four PCs and that for spectral PCA uses five PCs; hence, more time is taken, as compared to DPCA, which uses only one PC.

C. Proposed HPCA Techniques

Here we propose two HPCA methods, i.e., spatial–spectral HPCA and spectral–spatial HPCA, that were demonstrated for MISR data in Figs. 6 and 7. In the spatial–spectral HPCA technique, we propose to apply PCA along the spatial dimension first, followed by the spectral dimension. In this case, PCA is applied to the information from the different angles for one particular wavelength (spatial fusion). This is then followed by performing PCA on the first PCs resulting from the previous step for each wavelength (spectral fusion). The first PC resulting from this hierarchical fusion is then used in classification as a representative of the assimilated data. Later, it will be shown that performing PCA in such a hierarchical way can result in better classification accuracy. However, in the spectral–spatial HPCA technique, we apply PCA along the spatial dimension first, followed by the spectral dimension. In this case, PCA is applied to the information from the different wavelengths (spectral fusion) for one camera angle, which is similar to the spectral PCA method. This is then followed by applying PCA on the first PCs that were obtained from the first step (spectral fusion). The first PC resulting from this hierarchical fusion is used in classification as a representative of the assimilated data. Similar classification results are obtained for both HPCA methods. Applying the Sobel edge detection technique over the first PC that was obtained from both fusion methods would result in identifying the sudden changes in the spatial domain that correspond to abrupt change in the spectral domain due to different regions of dust intensity [11].

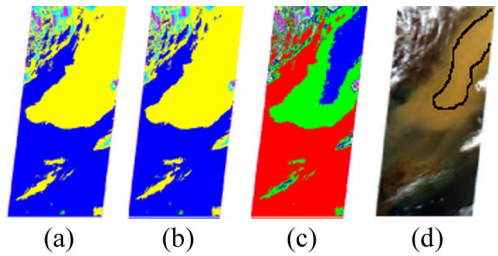

Fig. 8. Experimental results showing the classification obtained from the (a) original method, (b) DPCA spectral and spatial fusion, (c) HPCA spatial–spectral and spectral–spatial fusion, and (d) true color with visual highlight.

III. Comparative Results

Here, we present a quantitative and qualitative analyses of the experimental results that were obtained from the preceding techniques. The computational complexities of the six fusion schemes are also evaluated and compared for their computational savings.

A. Experimental Results

To compare our proposed HPCA techniques with the existing ones, the "percent variation" is used as a measure of information content in the fused images. Furthermore, we are using the mean squared error (MSE) and standard deviation (Stdev) to quantitatively assess the quality of fused information. This is done by comparing the obtained results using all original data and outputs that were received after performing the previously shown PCA.

K-means clustering is used to quantitatively examine the existing and proposed PCA-based fusion techniques aside from the direct classification, as described in Section II. The fused outputs are compared to the original data using the same threshold and the maximum number of iterations. In this method, since the original data have all the information, the classification results are assumed to have the highest classification accuracy [12]. K-means use spatial and spectral information from the image under study. It creates clusters of discrete classes having sets of similar objects. Good clustering results when the objects in the same class are more or less alike and the objects in different classes are, in some sense, different [12]. We determine the number of classes by adding one class at a time and then performing PCA until we reach a number after which the number of classes remained constant. In performing clustering of the results that were obtained from the original and the non-HPCA methods, similar classification results were observed, as shown in Fig. 8(b). This result demonstrates that PCA classification can at least achieve the same accuracy levels as the original data, using a much smaller data set. The classification results of the HPCA approach are shown in Fig. 8(c). Performing HPCA, by applying PCA spatially, followed by applying it spectrally and vice versa, has produced much better results. Specifically, it added a new class, which helps in distinguishing different regions of strength within the dust event, revealing strong and weak areas of dust concentration. In Fig. 8(d), a true color view of the dust event, which is the region of higher dust concentration, is highlighted by an expert Earth scientist. This area matches the obtained results from HPCA very closely.

TABLE I
(a) PERCENT VARIATION IN FIRST PC OBTAINED FROM DPCA. (b) PERCENT VARIATION IN FIRST PC OBTAINED FROM SPECTRAL FUSION. (c) PERCENT VARIATION IN FIRST PC OBTAINED FROM SPATIAL FUSION. (d) PERCENT VARIATION OF FIRST PC OBTAINED FROM HPCA

(a)

Camera	Percent Variation
DPCA	89.955

(b)

Camera	Percent Variation
NADIR (PC1)	99.695
Camera A (PC1)	99.321
Camera B (PC1)	99.093
Camera C (PC1)	98.987
Camera D (PC1)	98.915

(c)

Camera	Percent Variation
Near IR (PC1)	94.255
Red (PC1)	94.156
Green (PC1)	94.424
Blue (PC1)	94.614

(d)

Camera	Percent Variation
Spatial – Spectral	96.231
Spectral – Spatial	93.074

Clustering the original data and the data resulting from DPCA revealed some limitations because the higher dimensionality that is associated with the original data complicates the clustering process [12]. In HPCA, PCA is applied over data that were obtained from the same domain. Thus, as PCA concentrates information, concentrating data and removing redundancies within a given domain work more efficiently. When performing PCA, if the output components are arranged in descending order of information content, then most of the information can be obtained from the first few components. Table I shows a percent variation of the obtained PCs. After DPCA, the first PC contains more than 89% of the total data variance, whereas the first PC resulting from HPCA contains more than 93% of data variance, as compared to the original data. Although data variance is expected to be more in the first PC, since it is a data-dependent output, image artifacts result in a lower variance but still more than 90% in almost all cases.

We also notice from Table I that, by performing HPCA, the amount of information in the first PC is more, as compared to DPCA. This is consistent with the earlier stated result that k-means classification has better accuracy when the first PC is obtained from spatial PCA, followed by spectral PCA (or vice versa), as compared to DPCA.

Visual identification of the different dust concentration regions is corroborated by applying the Sobel filter over the obtained PC1 from both methods for quantitative validation of the observed regions of higher concentration. The different levels of dust concentrations based on the Sobel operator are clearly distinguished in the color-coded representation of Fig. 9(a).

This is because Sobel detects abrupt changes, among which are the edges of different dust concentration levels that are considered to be high-frequency events. Such abrupt changes in the spatial domain reflect high-frequency activity in the spectral domain; hence, the Sobel output from both methods, namely, frequency-based and camera-based, appear to be the same.

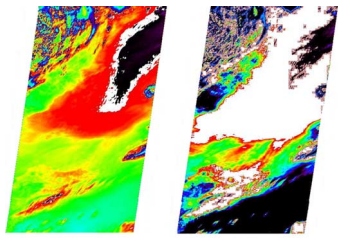

Fig. 9. (a) Sobel edge applied to PC1 showing different regions of dust concentration within the dust cloud. (b) Difference image between Sobel and k-means classification showing dust level missed from the classified image.

TABLE II
ACCURACY MEASURES BETWEEN THE ORIGINAL VERSUS REDUCED MISR DATA

Fusion Experiment	Type	MSE	Stdev
DPCA	PC1	0.1363	0.1159
	PC 1,2,3,4	0.0048	0.0689
Spectral PCA	PC 1,2,3,4,5	0.0019	0.0444
Spatial PCA	PC1,2,3,4	0.0126	0.1114
Spectral- Spatial PCA	PC of PC1's Spatial Fusion	0.0836	0.2768
Spectral- Spatial PCA	PC 1 of PC1's Spectral Fusion	0.0655	0.1835

Comparing Figs. 8(c), (d) and Fig. 9(a), a clear correspondence is observed between the Sobel, k-means classified, and the visual images in the general appearance of the dust event. However, the Sobel image revealed additional levels of dust accuracy, showing higher degrees of dust concentration, as compared to the k-means classified image. Hence, the difference image between the Sobel output and the classification output revealed the region of the dust cloud that was not identified through classification, as shown in Fig. 9(b).

This output manifests the fact that PCA has concentrated the information from both the spectral and the spatial domains in a way to better represent different regions of the dust event. Observing these dust regions, however, would require good mapping of the fused information. MSE measures the similarity between the original spectral signature and that of the PCA reduced image after data fusion to show the quality of image fusion. MSE values vary from 0 to 1.

Table II shows normalized MSE and Stdev representing the accuracy of the similarity between the original and reduced images for each of the aforementioned experiment. It is important to note that, as the number of PCs increases, MSE decreases, which means that the signal becomes more similar to the original data. A proportionate increase in Stdev can be observed with increase in the number of PCs. We also notice that the MSE that was obtained from DPCA is larger than that obtained from HPCA, which confirms that HPCA has more information than DPCA and, hence, produces better classification accuracy.

B. Computational Complexity

To quantitatively evaluate the effectiveness of the proposed techniques, we have used the k-means algorithm. As the dimension increases, clustering took more time, and this dependence is linear for the brute force method [13]. Dependence on dimension seems to be characteristic of many algorithms that are based on k-dimensional trees and many other variants,

TABLE III
TIME EXPRESSION OF DIFFERENT PCA METHODS

Fusion Experiment	Time Expression	Computational Complexity
K-Means	*(dkn)*	O(dkN)
DPCA	(MN^2+N^3)	$O(N^3)$
HPCA	$((N_2+ N_1)M + (N_2^2+ N_1^2)M+(N_2^3+ N_1^3))$	$O(N^2)$

such as R-trees. This dependence can also become exponential, depending on the filter used. For simplicity, we have used the brute force method, which is the best case scenario for this problem, as we are only considering the linear dependence on dimensionality. The computational complexity for the k-means algorithm is given as $O(dkn)$ [13], where d is the dimension of the data set to be classified, k is the number of classes in the data, and n is the number of points in the data set.

To fully evaluate the performance of the proposed methods, we have compared the computational efficiency of DPCA to HPCA. HPCA is computed in steps, and the resultant data are represented in the new uncorrelated coordinate system [12]. The weighting coefficient is obtained from the transposed matrix of eigenvectors of $\sum x$, which is the covariance matrix of x, where x represents the original pixel points. The computational complexity of the PCA algorithm over an M-pixel image of N spectral bands can be computed in four steps.

1) Find the mean vector: $O(MN)$.
2) Assemble the covariance matrix: $O(MN^2)$.
3) Perform eigenanalysis, i.e., generate the transformation matrix to compute the eigenvectors with a Jacobi method: $O(N^3)$.
4) Perform pixel-by-pixel linear transformation: $O(RMN)$.

Therefore, the total estimated time complexity of a PCA is $O(MN^2 + N^3)$, where M is the number of pixels of the image data, N is the number of bands, and R is the number of formed components $(R \leq N)$. Similarly, for computing PCA by parts, the estimated time complexity for the first is given as $O((R_1 + 1)M_1N_1 + M_1N_1^2 + N_1^3)$, where M_1 is the number of pixels of the image data, N_1 is the number of bands, and R_1 is the number of formed components $(R_1 \ll N_1)$. Stage two-time complexity can be calculated as shown here.

- Mean vector for second set: $O(M_1R_1N_2)$.
- Assemble covariance matrix: $O(M_1(R_1N_2)^2)$.
- Eigen analysis to generate the transformation matrix: $O((R_1N_2)^3)$.
- Linear transformation to obtain components: $(R_2M_1R_1N_2))$.

Here, M_1 is the number of pixels of the image data, (R_1N_2) is the number of bands for the second PCA, and R_2 is the number of obtained components $(R_2 \ll N_1R_2)$. The overall complexity of PCA is given as $O(M_1R_1N_2 + M_1(R_1N_2)^2 + (R_1N_2)^3 + R_2M_1R_1N_2)$. The overall complexity for this experiment is given as $O((R_1 + 1)M_1N_1 + M_1N_1^2 + N_1^3) + O((R_2 + 1)M_1R_1N_2 + M_1(R_1N_2)^2 + (R_1N_2)^3)$. Since, in our case, we are using $R_1 = R_2 = 1$, the overall complexity equation upon simplification becomes $O((N_2 + N_1)M_1 + (N_2^2 + N_1^2)M_1 + (N_2^3 + N_1^3))$. This shows that the time complexity of the HPCA method becomes additive. The obtained results are summarized in Table III.

IV. CONCLUSION AND FUTURE WORK

In this letter, we have presented efficient image fusion techniques for MISR data based on the hierarchical application of PCA both analytically and experimentally. The hierarchical application of PCA in the spatial domain for one spectral band across all the cameras, followed by the application of PCA across the top components from this step, has proven to be very useful. Likewise, the hierarchical application of PCA across all bands in each camera, followed by the application of PCA across the top components from that step, has also been shown to be very useful. Those hierarchical techniques seem to do better in classification, as they seem to be more capable of concentrating the related information from one domain much more easily. Accuracy measures were applied through the MSE and Stdev computation to compare the results that were obtained from different applied fusion methods with the original data. These hierarchical methods did not only produce improved classification accuracy but were also shown to be more computationally efficient being additive, as compared to being multiplicative for DPCA, as demonstrated by the computational complexity analyses. From the preceding discussion, it is clear that our method is innovative for time efficiency and for dust level discrimination. In our future work, we will investigate the impact of different backgrounds and different classification techniques on dust storm detection for dust events of different origins and geographic locations.

REFERENCES

[1] C. Pohl and J. L. Van-Genderen, "Multisensor image fusion in remote sensing: Concepts, methods and applications," *Int. J. Remote Sens.*, vol. 19, no. 5, pp. 823–854, Mar. 1998.
[2] S. Mukhopadhyay and B. Chanda, "Fusion of 2D grayscale images using multiscale morphology," *Pattern Recognit.*, vol. 34, no. 10, pp. 1939–1949, Oct. 2001.
[3] A. Solberg, T. Taxt, and A. Jain, "A Markov random field model for classification of multisource satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 1, pp. 100–113, Jan. 1996.
[4] Z. Zhang, S. Suan, and F. Zheng, "Image fusion based on median filters and SOFM neural networks: A three-step scheme," *Signal Process.*, vol. 81, no. 6, pp. 1325–1330, Jun. 2001.
[5] H. M. El-Askary, A. Agarwal, T. El-Ghazawi, M. Kafatos, and J. LeMoigne, "Enhancing dust storm detection using PCA based data fusion," in *Proc. IGARSS*, Jul. 25–29, 2005, vol. 2, pp. 1424–1427.
[6] M. Kafatos, H. M. El-Askary, and G. Kallos, "Potential for early warning of dust storms," in *Proc. Int. Symp. Sand and Dust Storm*, Beijing, China, Sep. 12–14, 2004.
[7] J. V. Martonchik, "Retrieval of surface directional reflectance properties using ground level multi-angle measurements," *Remote Sens. Environ.*, vol. 50, no. 3, pp. 303–316, 1994.
[8] H. M. El-Askary, R. Gautam, R. P. Singh, and M. Kafatos, "Dust storms detection over the Indo-Gangetic basin using multi sensor data," *Adv. Space Res. J.*, vol. 37, no. 4, pp. 728–733, 2006.
[9] L. Wald, "A European proposal for terms of reference in data fusion," *Int. Arch. Photogramm. Remote Sens.*, vol. XXXII, pt. 7, pp. 651–654, 1998.
[10] L. Wald, "Some terms of reference in data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1190–1193, May 1999.
[11] H. M. El-Askary, S. Sarkar, M. Kafatos, and T. El-Ghazawi, "A multisensor approach to dust storm monitoring over the Nile Delta," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 10, pp. 2386–2391, Oct. 2003.
[12] J. A. Richards and X. Jia, "Segmented principal components transformation for efficient hyperspectral remote sensing image display and classification," *IEEE Trans.Geosci. Remote Sens.*, vol. 37, no. 1, pp. 538–542, Jan. 1999.
[13] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 881–892, Jul. 2002.