Spectral/Spatial Hyperspectral Image Compression in Conjunction with Virtual Dimensionality

Bharath Ramakrishna[1] Jing Wang[1] Chein-I Chang[1] Antonio Plaza[1,2] Hsuan Ren[3]
Chein-Chi Chang[4] Janet L. Jensen[5] James O. Jensen[5]

[1]Remote Sensing Signal and Image Processing Laboratory
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County, Baltimore, MD 21250
[2]Computer Science Department, University of Extremadura
Avda. de la Universidad s/n,10.071 Caceres, SPAIN
[3]Center for Space and Remote Sensing Research
Department of Information Engineering
National Central University, Chungli, Taiwan, ROC
[4]Department of Civil and Environmental Engineering
University of Maryland, Baltimore County, Baltimore, MD 21250
[5]US Army Edgewood Chemical and Biological Center
Aberdeen Proving Ground, MD 21010

Abstract

Hyperspectral image compression can be performed by either 3-D compression or spectral/spatial compression. It has been demonstrated that due to high spectral resolution hyperspectral image compression can be more effective if compression is carried out spectrally and spatially in two separate stages. One commonly used spectral/spatial compression implements principal components analysis (PCA) or wavelet for spectral compression followed by a 2-D/3D compression technique for spatial compression. This paper presents another type of spectral/spatial compression technique, which uses Hyvarinen and Oja's Fast independent component analysis (FastICA) to perform spectral compression, while JPEG2000 is used for 2-D/3-D spatial compression. In order to determine how many independent components are required, a newly developed concept, virtual dimensionality (VD) is used. Since the VD is determined by the false alarm probability rather than the commonly used signal-to-noise ratio or mean squared error (MSE), our proposed FastICA-based spectral/spatial compression is more effective than PCA-based or wavelet-based spectral/spatial compression in data exploitation.

1. INTRODUCTION

Hyperspectral image compression has received considerable interest in recent years because hyperspectral imaging sensors have become very popular and the acquired hyperspectral image data are enormous and highly correlated. Many algorithms have been developed for hyperspectral image compression in the past. Two approaches are generally taken. One is to consider a hyperspectral image as an image cube and then apply 3-D compression which is directly extended from 2-D compression. Most notable techniques arising from the approach of this type are 3D-JPEG2000 [1] which is an extension of JPEG2000 [2] and 3D set partitioning in hierarchical trees (SPIHT) [3] which is a result of Said-Pearlman's 2D set partitioning in hierarchical trees (SPIHT) [4]. A second approach is to perform compression spectrally and spatially. More specifically, instead of performing 3D compression on an image cube as the first approach does, the spectral/spatial compression performs a two-stage compression process, spectral compression in the first stage followed by spatial compression in the second stage [5-6]. In the first stage of spectral compression, principal components analysis (PCA) is generally used to de-correlate spectral information provided by hundreds of spectral bands. The resulting spectral de-correlated image cube is then compressed by either 3D or 2D compression to achieve spatial compression. Surprisingly, as shown in [6] a simple PCA-based spectral/2D spatial compression can perform at least as

Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI,
edited by Sylvia S. Shen, Paul E. Lewis, Proceedings of SPIE Vol. 5806
(SPIE, Bellingham, WA, 2005) · 0277-786X/05/$15 · doi: 10.1117/12.604128

well as 3D compression. This suggests that spectral/spatial compression may be more appropriate and effective than 3D compression in hyperspectral image compression.

Spectral/spatial lossy compression is not new and has been studied extensively in data compression [5]. The basic idea is to use the PCA to perform DR to achieve spectral compression. A key issue in this spectral compression is the number of principal components (PCs) needed to be retained during lossy compression. In the past, this issue has been addressed by preserving a certain percentage of energy in terms of accumulated sum of eigenvalues. Unfortunately, it was shown in [6] that was not an appropriate criterion for determination of number of PCs. Instead, a new concept, called virtual dimensionality (VD) introduced in [7-8] was shown to be more effective than the use of percentage of energy [6]. The work reported in [6] was the first which ever used the VD to determine number of PCs required to be retained during DR for spectral compression. However, an underlying assumption of using the PCA for DR is based on the fact that the data are well-represented and structured in terms of variance, where most of data points are clustered and can be packed in a low dimensional space. Unfortunately, it was recently shown in [9-10] that signal-to-noise ratio (SNR) is a better measure than data variance to measure image quality in multispectral imagery. Similarly, the mean squared error (MSE) has been also widely used as a criterion for optimality in communications and signal processing such as quantization. But, it is also known that it may not be appropriate to be used as a measure of image interpretation. This is particularly true for hyperspectral imagery which can uncover many unknown signal sources, some of which may be very important in data analysis such as anomalies, small targets which generally contribute very little to SNR or MSE. In the PCA these targets may only be retained in minor components instead of principal components. So, preserving only the first few principal components may lose these targets. In SNR or MSE, such targets may very likely be suppressed by lossy compression if no extra caution is taken since missing these targets may only cause inappreciable loss of signal energy or small error. In order to resolve this dilemma, high-order statistics-based dimensionality reduction techniques must be sought. One such approach is independent component analysis (ICA) [11] which makes use of statistical independency as criterion to separate components rather than second-order de-correlation performed by the PCA. In this paper, we will investigate ICA-based spectral/spatial compression for hyperspectral images. More specifically, spectral compression will be performed by ICA for dimensionality reduction (DR), while the spatial compression will be performed by either 2D/3D-JPEG-2000 or 2D/3D-SPHIT.

One major problem with replacing the PCA by the ICA for spectral compression is that there is no criterion which can be used to prioritize the components produced by the ICA as does the PCA which uses the variance to prioritize its produced principal components. As a result, the ICA generally cannot perform dimensionality reduction as the way the PCA does. Additionally, the ICA-generated components are statistically independent, not merely second-order de-correlated. Therefore, using the ICA for spectral compression is not a simple replacement of the PCA with the ICA and may not be as easy as it was thought. Two major issues need to be addressed, which are determination of the number of components and prioritization of components for the ICA. The first issue can be dealt with by the VD in the same manner that the VD is used to determine the number of PCs for the PCA for dimensionality reduction. Unfortunately, the second issue only exists in the ICA and never becomes a problem for the PCA since the PCA always prioritizes its PC's in accordance with the magnitude of eigenvalues. The cause of this issue is due to the fact that the components produced by the ICA are generated by using random projection vectors as initial conditions. Since the initial projection vectors are randomly generated, the appearing order of independent components (ICs) generated by the ICA is not consistent every time the ICA is implemented. This is completely different from the PCA-generated PCs where the first PC is always the one specified by the largest eigenvalue, the second PCs is specified by the second largest eigenvalue and so on. In this case, once the VD determines the number of dimensions, say p, to be retained the first p PCs that correspond to the first largest p eigenvalues will be used for dimensionality reduction. This is certainly not true for ICA-generated ICs. That is, the same IC is not necessary to appear in the same order all the time. On the other hand, an earlier appearance of an IC does not necessarily imply that it is more significant or important than one appearing later. As a result, the first p ICs will not be the same ones if the ICA is implemented twice. Therefore, in this case, when the ICA performs DR, we generally do not know which ICs must be retained and which ones must be left out. Recently, the issue of using the ICA for DR was investigated in [12] where two algorithms were developed to prioritize the ICs. This paper explores ICA-based DR algorithms in application of hyperspectral image compression and further compares the ICA-based spectral/spatial compression to PCA-based spectral/spatial compression via experiments for performance analysis.

2. ICA

Independent component analysis (ICA) has emerged as one of major signal processing techniques in recent years. It has versatile applications ranging from channel equalization, blind source separation, speech analysis to functional magnetic

resonance imaging (fMRI) [11]. Its applications to linear mixture analysis for hyperspectral images have been also investigated in [14-16]. The key idea of the ICA is to use a linear mixture model to demix a set of separate independent sources. In order to validate its approach, an underlying assumption is that at most one source in the mixture model can be allowed to be a Gaussian source. This is due to the fact that a linear mixture of Gaussian sources is still a Gaussian source. More precisely, let **x** be a mixed signal source vector expressed by

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{1}$$

where **A** is a mixing matrix and **s** is a signal source vector to be separated. The purpose of the ICA is to find a demixing matrix **W** that separates the signal source vector **s** into individual sources which are statistically independent. Several different criteria have been proposed to measure source independency [11]. Nevertheless, they all originated from the concept of mutual information which is a criterion to measure the discrepancy between two random sources [17].

3. ICA-BASED DIMENSIONALITY REDUCTION

Over the past years, DR is generally performed by the PCA. Interestingly, to the authors' best knowledge, there is little work of applying ICA to DR reported in the literature. One possible reason is that the ICA was not originally developed for the purpose of DR. Also, unlike the PCA which prioritizes its generated principal components in accordance with magnitude of eigenvalues, there is no specific criterion to rank components produced by the ICA. Since the ICA is a well-established technique, we will only focus on the issues that arise in DR.

In order to implement the ICA, the FastICA algorithm developed by Hyvarinen and Oja [13] was used to find ICs. For each spectral band image, it was converted to a vector. More specifically, assume that a hyperspectral image cube has size of $M \times N \times L$ where L is the number of spectral bands and MN is the size of each spectral band image. The hyperspectral image cube can then be represented by a data matrix **X** of size $L \times MN$ with L rows and MN columns. In other words, each row in the data matrix **X** is specified by a particular spectral band image. As a result, a total of L ICs can be generated by the FastICA. However, as noted r=Mα, a hyperspectral image pixel can be generally considered as a linear mixture of a set of known image endmembers where the number of endmembers, p is generally much smaller than L, that of spectral bands. In this case, when the DR is performed, only p ICs are required and there is no need of producing all ICs for image analysis. But, it also gives rise to an issue on that which p ICs must be selected for DR.

According to the introduction, there is a need for ICA-DR to address problems that cannot be resolved by either the PCA-DR or MNF-DR. However, there are some major issues to implement ICA for DR (in our case, the FastICA for DR). First of all, the FastICA-generated ICs are not necessarily in order of information significance as the way that PCs are generated by the PCA or the MNF in accordance with decreasing magnitude of eigenvalues or SNRs. Another is that ICs generated by the FastICA in different runs do not necessarily appear in the same order. These issues are primarily due to the nature that the initial projection unit vectors used to produce ICs by the FastICA are randomly generated. Therefore, an IC generated earlier by the FastICA is not necessarily more significant than one generated later. In order to resolve the issue in the use of the random initial projection unit vectors, a FastICA-based algorithm, called ICA-DR is developed for DR. The proposed ICA-DR considers each generated IC as a random variable. In light of this interpretation, the i-th IC_i can be assumed and described by a random variable ζ_i with values specified by the gray level value of the n-th pixel in the IC_i, denoted by z_n^i. In this case, the FastICA-generated ICs can be ranked and prioritized by high-order statistics-based criteria. The idea of ICA-DR is to first determine the number of ICs needed to be retained, p that can be estimated by the VD. It then prioritizes the FastICA-generated ICs according to a high-order statistics criterion to select the first p prioritized ICs. The detailed implementation of the ICA-DR is summarized as follows.

ICA-DR Algorithm

1 Use the VD to determine the number of dimensions, p, required to be retained.
2 Use the FastICA to find $2p$ independent components, $\{\mathrm{IC}_i\}_{i=1}^{2p}$. It should be noted that for each IC the FastICA randomly generates a unit vector as an initial projection vector to produce the final desired projection vector for that particular component.
3 Calculate the following criterion for IC_i that combines 3rd and 4th orders of statistics for ζ_i.

$$J(\mathrm{IC}_i) = (1/12)\left[\kappa_i^3\right]^2 + (1/48)\left[\kappa_i^4 - 3\right]^2 \tag{2}$$

where $\kappa_i^3 = E\left[\zeta_i^3\right] = \frac{1}{MN}\sum_{n=1}^{MN}\left(z_n^i\right)^3$ and $\kappa_i^4 = E\left[\zeta_i^4\right] = \frac{1}{MN}\sum_{n=1}^{MN}\left(z_n^i\right)^4$ are sample means of 3rd and 4th orders of statistics in the IC_i.

4 Prioritize the $\{\mathrm{IC}_i\}_{i=1}^{2p}$ in accordance with the magnitude of $J(\mathrm{IC}_i)$.

5. Select those ICs with the first p largest $J(\mathrm{IC}_i)$ to perform DR.

It should be noted that ICA-DR is supposed to run and prioritize all the ICs, then selects the first p prioritized ICs. However, in practice this is not necessary. According to our experiments, the VD can be used to set an upper bound on the number of ICs required to be generated by the FastICA without having the FastICA run through all ICs. A good upper bound is empirically shown to be twice the VD to avoid small targets being left out.

4. ICA-BASED SPECTRAL/SPATIAL COMPRESSION

Depending upon whether or not an image reconstruction is required, there are two ways to implement the ICA-based spectral/spatial compression depicted in Figs. 1-2 respectively. In Fig. 1, when only IC images are of major interest, the compression is resulting from the reconstructed IC images, referred to as compressed domain.

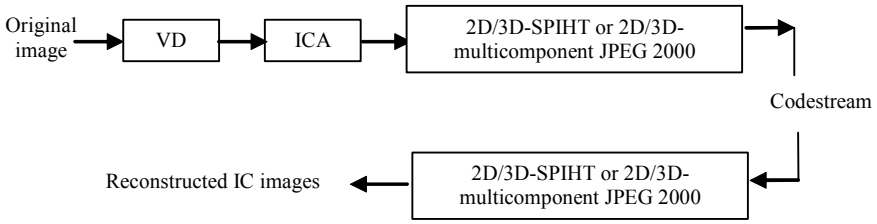


Figure 1. ICA-based spectral/spatial compression without image reconstruction

On the other hand, when an image reconstruction is required, an inverse ICA (IICA) is then applied to reconstruct the original image based on IC images as shown in Fig. 2. In this case, the compression is resulting from the reconstructed image cube, referred to as de-compressed image.

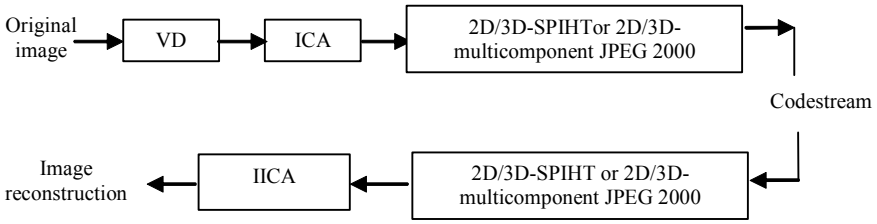
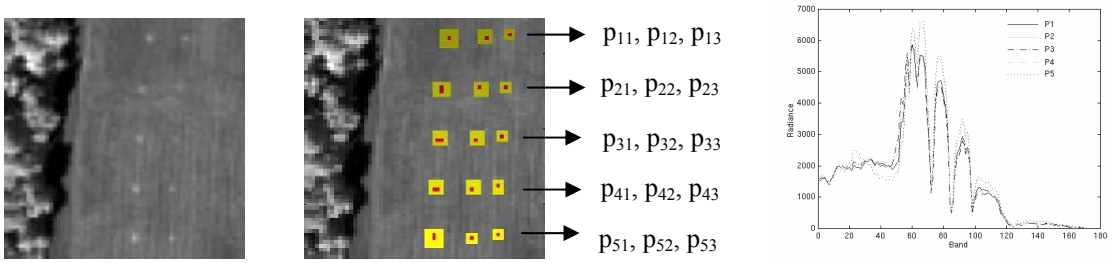


Figure 2. ICA-based spectral/spatial compression for image reconstruction

5. EXPERIMENTS

In this section, real HYperspectral Digital Imagery Collection Experiments (HYDICE) image is used to demonstrate that a direct application of 3-D data compression to a hyperspectral image without extra care may result in significant loss of information. It is an image scene shown in Fig. 3(a), which has a size of 64×64 pixel vectors with 15 panels in the scene and the ground truth map in Fig. 3(b). It was acquired by 210 spectral bands with a spectral coverage from 0.4μm to 2.5μm. Low signal/high noise bands: bands 1-3 and bands 202-210; and water vapor absorption bands: bands 101-112 and bands 137-153 were removed. So, a total of 169 bands were used. The spatial resolution is 1.56m and the spectral resolution is 10nm. Within the scene in Fig. 3(a) there is a large grass field background, and a forest on the left edge. Each element in this matrix is a square panel and denoted by p_{ij} with rows indexed by i and columns indexed by j. For each row $i = 1,2,\ldots,5$, there are three panels p_{i1}, p_{i2}, p_{i3}, painted by the same material but with three different sizes. For each column $j = 1,2,3$, the 5 panels p_{1j}, p_{2j}, p_{3j}, p_{4j}, p_{5j} have the same size but with five different materials. Nevertheless, they were still considered as different materials. The sizes of the panels in the first, second and third

columns are $3\text{m} \times 3\text{m}$, $2\text{m} \times 2\text{m}$ and $1\text{m} \times 1\text{m}$ respectively. Since the size of the panels in the third column is $1\text{m} \times 1\text{m}$, they cannot be seen visually from Fig. 3(a) due to the fact that its size is less than the 1.56m spatial resolution. Fig. 3(b) shows the precise spatial locations of these 15 panels where red pixels (R pixels) are the panel center pixels and the pixels in yellow (Y pixels) are panel pixels mixed with the background. The 1.56m-spatial resolution of the image scene suggests that most of the 15 panels are one pixel in size except that p_{21}, p_{31}, p_{41}, p_{51} which are two-pixel panels. Fig. 3(c) plots the 5 panel spectral signatures $\{\mathbf{P}_i\}_{i=1}^{5}$ with $\mathbf{P}_i$ obtained by averaging R pixels in the $3\text{m} \times 3\text{m}$ and $2\text{m} \times 2\text{m}$ panels in row *i* in Fig. 3(b).



(a) (b) (c)

Figure 3. (a) A HYDICE panel scene which contains 15 panels; (b) Ground truth map of spatial locations of the 15 panels; (c) five panel signatures $\{\mathbf{P}_i\}_{i=1}^{5}$

It should be noted the average of R pixels in the 1m x 1m, 2m x 2m and 3m x 3m were used as required prior target knowledge for the following comparative studies. Table 1 gives the VD estimates for the HYDICE image in Fig. 3(a) where VD = 9 was chosen for our experiments.

Table 1. VD estimates for the HYDICE scene in Fig. 1 with various false alarm probabilities

	$P_F = 10^{-1}$	$P_F = 10^{-2}$	$P_F = 10^{-3}$	$P_F = 10^{-4}$	$P_F = 10^{-5}$
VD	14	11	9	9	7

Five lossy compression techniques, (1) 3D-multicomponent JPEG2000, (2) PCA/3D-multicomponent-JPEG2000, (3) IPCA/3D-multicomponent JPEG2000, (4) ICA/3D-multicomponent-JPEG2000 and (5) IICA/3D-multicomponent-JPEG2000, were selected for performance comparison with compression ratio (CR) = 20:1, 40:1, 80:1 and 160:1. Such selection was based on following two reasons. First of all, it has been shown in [6,19] that both JPEG2000 and SPIHT performed comparably. Therefore, only the JPEG2000 results are presented here. We refer details of SPHIT compression to [6,19]. Additionally, it has been also shown in [6,19] that 2D spatial compression did not perform as well as 3D spatial compression in spectral/spatial compression. So, only 3D spatial compression, 3D-multicomponent JPEG2000 was used to perform spatial compression after spectral compression. In order to see how these five lossy compression affects panel detection and classification, the constrained energy minimization (CEM) [8,18] was used for performance evaluation due to its effectiveness in detection and classification of the 15 panels in [8]. Fig. 4 shows the 15-panel detection and classification of the CEM based on the original un-compressed image in Fig. 3(a) where the five panel signature $\{\mathbf{P}_i\}_{i=1}^{5}$ obtained in Fig. 3(c) were used as desired signatures for 15 panel detection and classification.

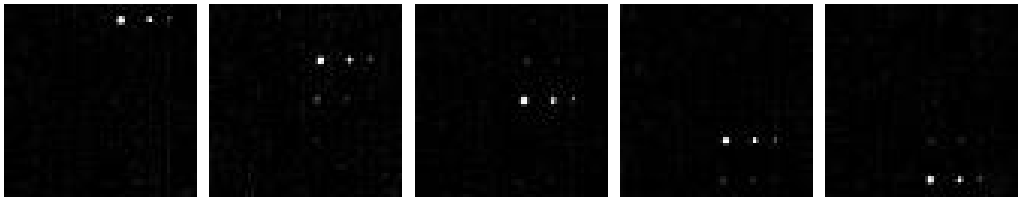
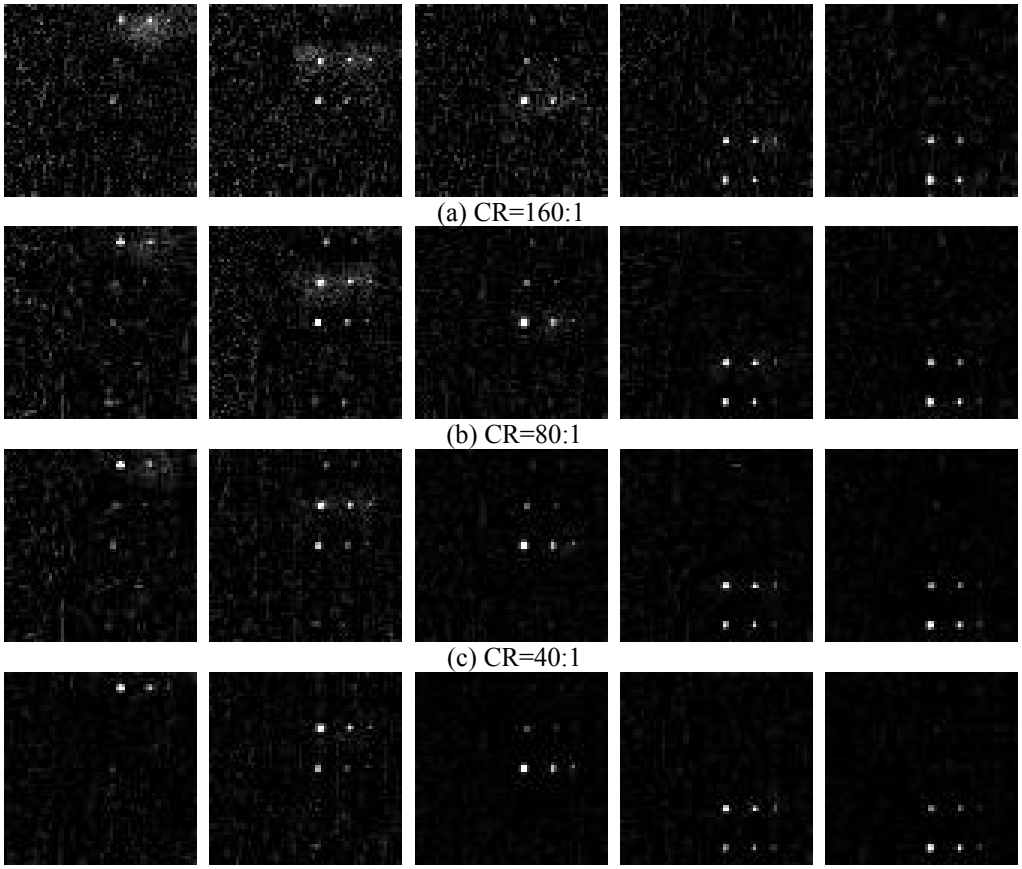


Figure 4. 15-panel detection and classification results of CEM operating on the original uncompressed image

Fig. 5 shows the panel detection results of the CEM operating on the 3D-JPEG2000 Multicomponent compressed image, which were obviously not as good as those in Fig. 4.



(a) CR=160:1

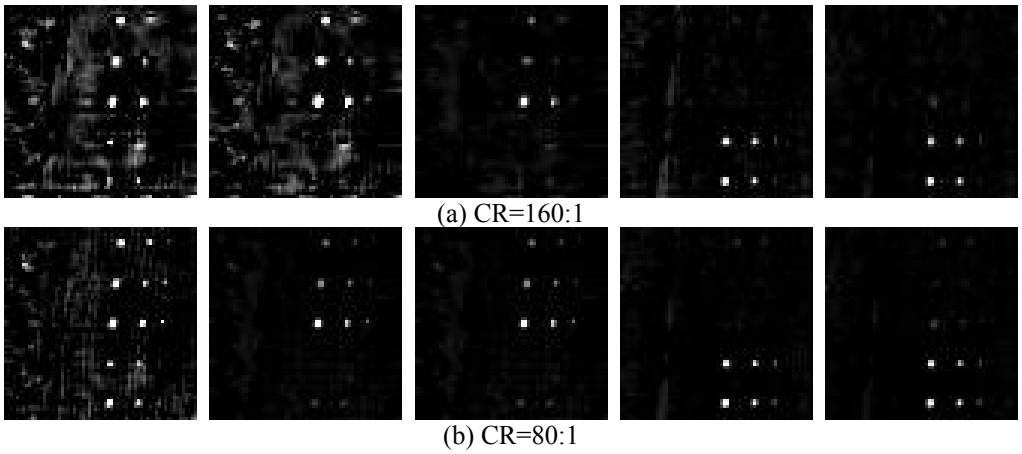
(b) CR=80:1

(c) CR=40:1

(d) CR=20:1

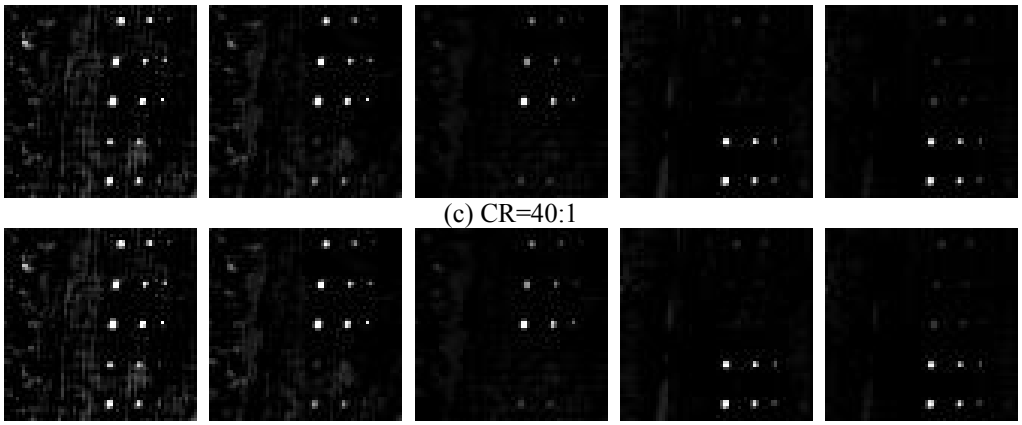
Figure 5. 15-panel detection and classification results of the 3D-JPEG2000 Multicomponent compressed image

If the CEM was applied to the PCA/3D-JPEG2000 Multicomponent compressed image, the results are shown in Fig. 6 and were even worse than those in Fig. 5.



(a) CR=160:1

(b) CR=80:1

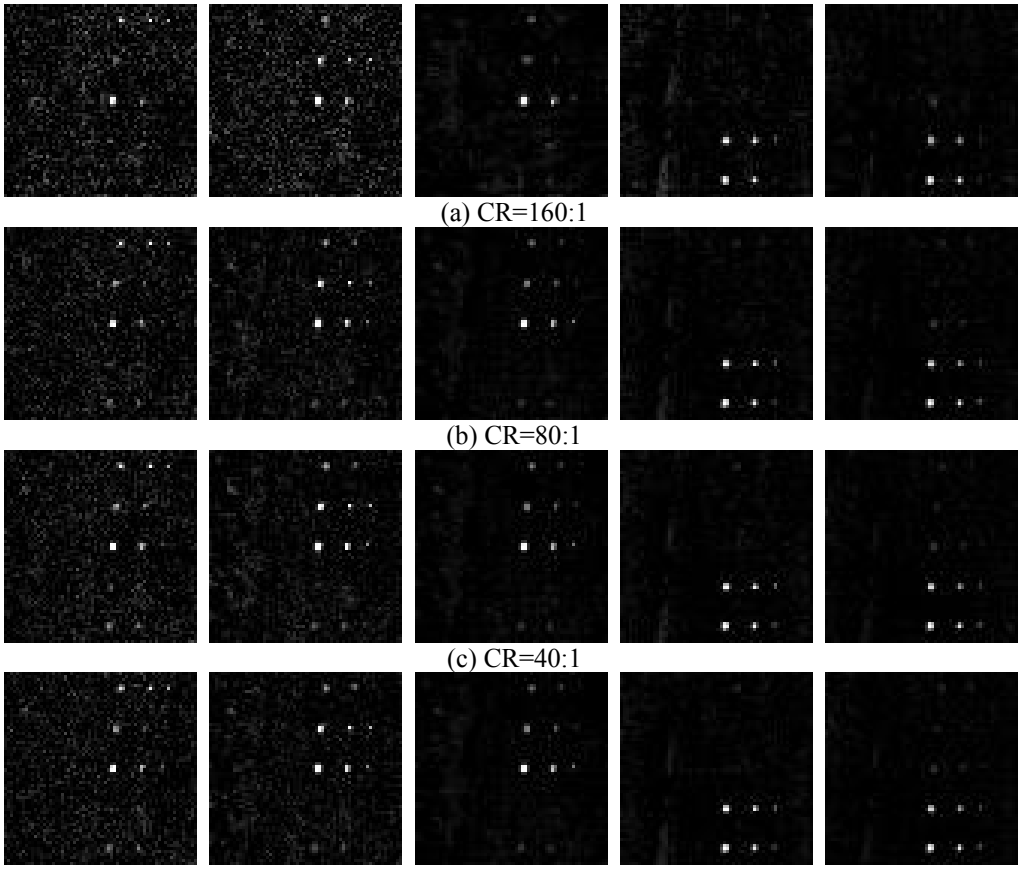


(c) CR=40:1

(d) CR=20:1

Figure 6. 15-panel detection and classification results of the PCA/3D-JPEG2000 Multicomponent compressed image

In order to see that the results in Fig. 6 could be improved if the reconstructed image was used instead of the compressed image, the IPCA/3D-JPEG2000 Multicomponent compressed image was implemented and results are shown in Fig. 7 where the CEM performance in panel detection and classification did improve slightly but not significantly.



(a) CR=160:1

(b) CR=80:1

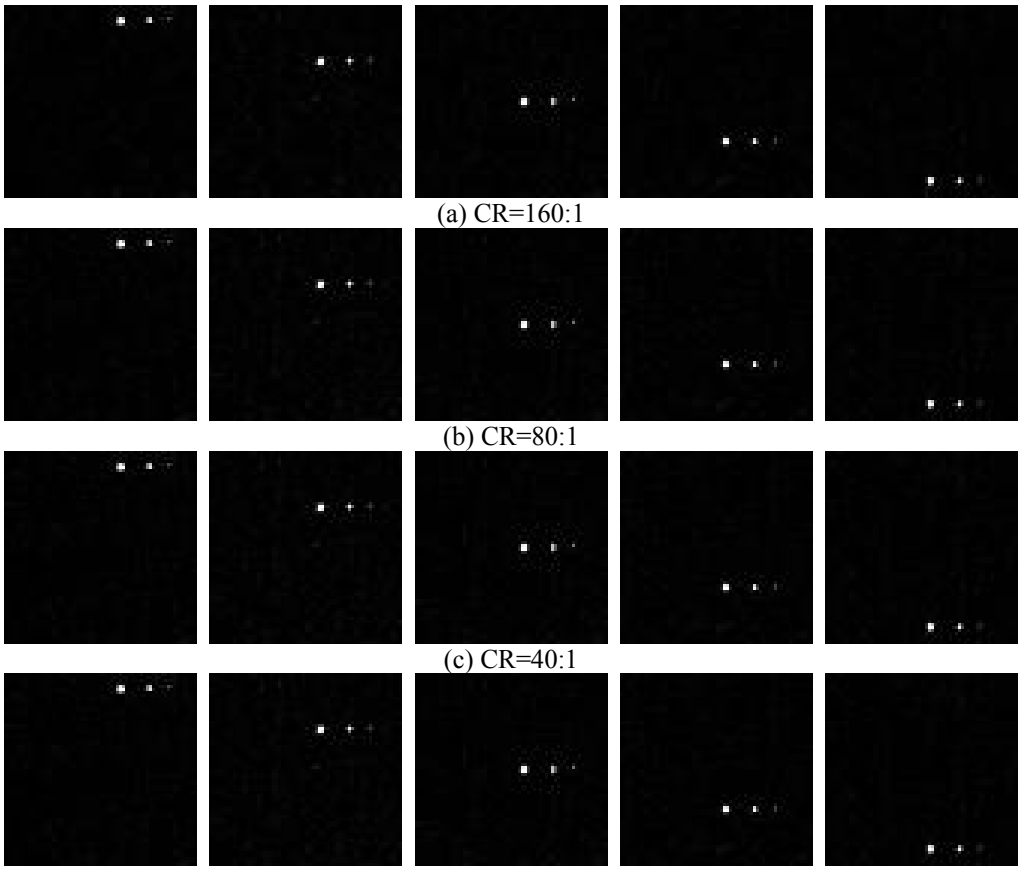
(c) CR=40:1

(d) CR=20:1

Figure 7. 15-panel detection and classification results of the IPCA/3D-JPEG2000 Multicomponent compressed image

Interestingly, when the PCA was replaced with the ICA in spectral/spatial lossy compression, the 15-panel detection and classification by the CEM shown in Fig. 8 not only improved substantially, but also was comparable to the results

obtained in Fig. 4 even in high compression ratio case such as 80:1, 160:1. Most surprisingly, according to Fig. 8, it seemed that the compression ratio did not play a role and had no factor at all. This was attributed to the effectiveness of the ICA-DR which preserved all the crucial panel information required for the CEM in detecting and classifying 15 panels.



(a) CR=160:1

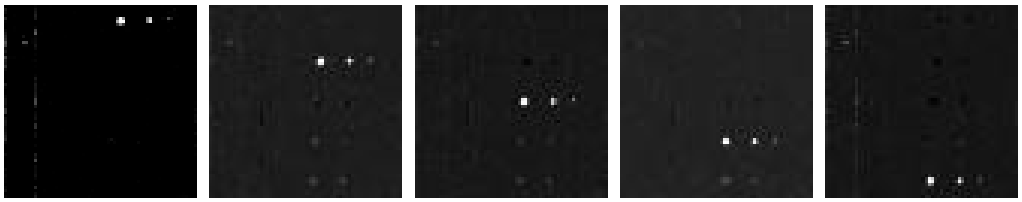
(b) CR=80:1

(c) CR=40:1

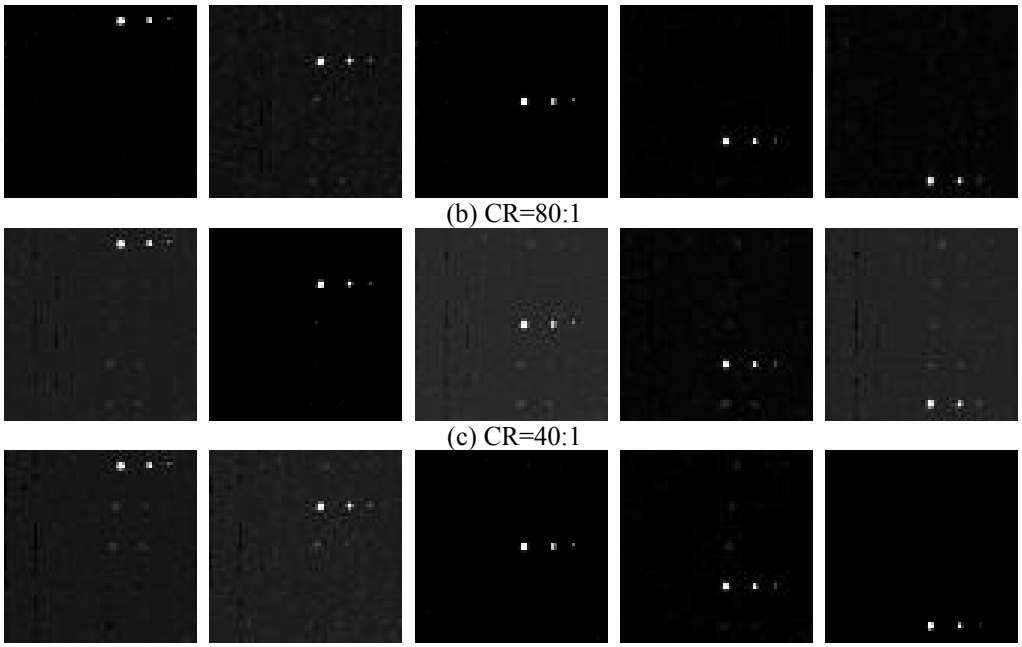
(d) CR=20:1

Figure 8. 15-panel detection and classification results of the ICA/3D-JPEG2000 Multicomponent compressed image

Analogous to the IPCA/spatial compression, the IICA/3D-JPEG2000 Multicomponent was also applied to see if it could improve on the performance of the ICA/3D-JPEG2000 Multicomponent. Fig. 9 shows the 15-panel detection and classification produced by the CEM where the results produced for CR = 20:1, 40:1 and 160:1 seemed degraded slightly. This may be due to the fact that the ICA-reconstructed image based on DR might introduce interfering information. This made sense because the ICA usually extracted subtle target information which may interfere performance.



(a) CR=160:1



(b) CR=80:1

(c) CR=40:1

(d) CR=20:1

Figure 9. Panel classification results of the IICA/3D-JPEG2000 Multicomponent compressed image

Finally, Table 2 tabulates SNR and MSE for each CR for Figs. 5-9 where ICA-based spectral/spatial compression produced the worst MSEs and SNRs but they achieved best classification results.

Table 2. SNR and MSE for each CR Figs.5-9

CR	160		80		40		20	
Method	SNR	MSE	SNR	MSE	SNR	MSE	SNR	MSE
3D (Fig. 5)	26.06	1.53E+06	29.80	6.49E+05	34.21	2.35E+05	39.60	6.79E+04
PCA/3D (Fig. 6)	14.28	1.67E+06	25.36	1.30E+05	46.71	9.55E+02	60.36	4.12E+01
IPCA/3D (Fig. 7)	36.07	1.53E+05	41.91	3.99E+04	42.79	3.26E+04	42.79	3.26E+04
ICA/3D(Fig. 8)	14.57	3.14E-01	23.50	4.02E-02	44.89	2.92E-04	50.62	7.80E-05
IICA/3D (Fig. 9)	11.60	4.28E+07	11.60	4.28E+07	11.61	4.28E+07	11.61	4.28E+07

The above table clearly shows that MSE and SNR are not appropriate criteria used for exploitation-based hyperspectral image compression, panel detection and classification in our experiments. Accordingly, using the MSE and SNR for hyperspectral image compression may mislead image analysts to an incorrect interpretation.

6. CONCLUSION

This paper presents a new approach to ICA-based spectral/spatial compression. It first performs the FastICA to reduce dimensionality while retaining desired information in reduced independent components. Then a 2D or 3D spatial compression technique is applied to either compressed independent component images or inverse ICA de-compressed images. Two major issues are investigated for this approach. One is the number of dimensions for reduction. This issue is resolved by the virtual dimensionality (VD) which provides a good estimate in determining how many components required to be retained for dimensionality reduction. Another issue to be resolved is how to rank ICA-generated components. The use of eigenvalues in ranking the principal components generated by the PCA is no longer valid for the ICA-generated components. In this case, new approaches must be sought. To address this problem, ICA-DR was introduced in this paper to prioritize ICA-generated components. Experiments demonstrate that the proposed ICA-based spectral/spatial compression in conjunction with the VD and component prioritization algorithms produce encouraging results and it generally performs better than the PCA-based spectral/spatial compression and 3D compression.

REFERENCES

[1] ISO, Information Technology—JPEG 2000 Image Coding System - Part 2: Extensions; Final Committee Draft, ISO, Geneva, Switzerland, Dec. 2000.

[2] D. S. Taubman and M. W. Marcellin, JPEG2000: Image Compression Fundamentals, Standard and Practice. Boston, MA: Kluwer, 2002.

[3] B.-J. Kim, Z. Xiong, and W.A.Pearlman, "Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT)," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 8, pp.1374-1387, December 2000.

[4] A. Said and W.A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. on Circuits and systems for Video Technology*, vol. 6, no. 3, pp. 243-350, June 1996.

[5] S. S. Shen and B. S. Beard, "Effects of hyperspectral compression on non-literal exploitation," *Proceedings of SPIE,* vol. 3438, pp. 191-199, 1998.

[6] B. Ramakishna, A. Plaza, C.-I Chang, H. Ren, Q. Du and C.-C. Chang, Spectral/Spatial Hyperspectral Image Compression, *Hyperspectral Data Compression*, edited by G. Motta and J. Storer, Springer-Verlag, 2005.

[7] C.-I Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 42, no. 3, pp. 608-619, March 2004.

[8] C.-I Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Kluwer Academic/Plenum Publishers, New York, N.Y., 2003

[9] A.A. Green, M. Berman, P. Switzer and M.D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 65-74, January 1988.

[10] J.B. Lee, A.S. Woodyatt and M. Berman, "Enhancement of high spectral resolution remote sensing data by a noise-adjusted principal components transform," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 28, no. 3, pp. 295-304, May 1990.

[11] A. Hyvarinen, J. Karhunen and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.

[12] J. Wang and C.-I Chang, "Dimensionality reduction by independent component analysis for hyperspectral image analysis," *IEEE International Geoscience and Remote Sensing Symposium*, Seoul, Korea, July 25-29, 2005.

[13] A. Hyvarinen and E. Oja, "A fast fixed-point for independent component analysis," *Neural Computation*, vol. 9, no. 7, pp. 1483-1492, 1997.

[14] J. Bayliss, J.A. Gualtieri and R.F. Cromp (1997), "Analyzing hyperspectral data with independent component analysis," *Poceedings of SPIE*, vol. 3240, pp. 133-143, 1997.

[15] T.M. Tu, "Unsupervised signature extraction and separation in hyperspectral images: a noise-adjusted fast independent component analysis approach," *Optical Engineering*, vol. 39, no 4, pp. 897-906, 2000.

[16] C.-I Chang, S.S. Chiang, J.A. Smith and I.W. Ginsberg, "Linear spectral random mixture analysis for hyperspectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 40, no. 2, pp. 375-392, February 2002.

[17] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, 1991.

[18] J.C. Harsanyi, *Detection and Classification of Subpixel Spectral Signatures in Hyperspectral Image Sequences*, Department of Electrical Engineering, University of Maryland, Baltimore County, MD, August 1993.

[19] Bharath Ramakrishna, *Principal Components Analysis (PCA)-Based Spectral/Spatial Hyperspectral Image Compression*, MS. Thesis, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD, 2004.