Multitemporal Hyperspectral Image Compression

Wei Zhu, *Student Member, IEEE*, Qian Du, *Senior Member, IEEE*, and James E. Fowler, *Senior Member, IEEE*

***Abstract*—The compression of multitemporal hyperspectral imagery is considered, wherein the encoder uses a reference image to effectuate temporal decorrelation for the coding of the current image. Both linear prediction and a spectral concatenation of images are explored to this end. Experimental results demonstrate that, when there are few changes between two images, the gain in rate-distortion performance is achieved over the independent coding of the current image. In addition, a strategy that explicitly removes salient temporal changes and stores them losslessly in the bitstream is proposed, and it is observed that this change-removal process results in a slight decrease in the rate-distortion performance with the benefit of perfect representation of the changed pixels.**

***Index Terms*—Data compression, hyperspectral imagery, multitemporal imagery.**

I. Introduction

OFTEN in hyperspectral applications, multiple images are collected for the same area at different times. From these multitemporal images, changes can be detected and analyzed. However, most prior literature on hyperspectral compression focuses on compression of only a single image. In the multitemporal setting though, temporal redundancy can be exploited, and rate-distortion performance improved when the encoder has access to a reference image that is highly correlated to the current image being coded. In this case, salient temporal changes between the reference and current images are special features that must be preserved well.

It has been shown that principal component (PC) analysis (PCA) in conjunction with JPEG2000 provides efficient rate-distortion performance for hyperspectral image compression, wherein PCA is applied for spectral decorrelation and JPEG2000 provides spatial coding of PC images. In particular, PCA outperforms a discrete wavelet transform (DWT) as a spectral transform for JPEG2000 [1]–[3]. In addition, explicitly eliminating minor PCs and compressing only major PCs can yield even better performance than a coding of the full PCA transform, particularly at low bit rates; this truncated spectral transform was called SubPCA in [4].

In this letter, we investigate the application of PCA and SubPCA for JPEG2000-based compression of multitemporal hyperspectral imagery. In particular, we consider several strategies in which an encoder exploits a reference image to effectuate temporal decorrelation for the coding of the current image; these include a linear prediction, as well as the concatenation of two images in the spectral direction. In addition, in order to preserve salient changes between the reference and current images, we develop a procedure to explicitly remove changes from the current image, losslessly storing them in the compressed bitstream.

The remainder of this letter is organized as follows. First, in Section II, we overview change detection for hyperspectral imagery. Then, in Section III, we explore several alternatives for temporal decorrelation in multitemporal imagery, as well as explicit removal of changes prior to compression. Finally, we present experimental results in Section IV and make some concluding remarks in Section V.

Manuscript received June 15, 2010; revised August 6, 2010; accepted September 22, 2010. Date of publication November 8, 2010; date of current version April 22, 2011.

The authors are with the Department of Electrical and Computer Engineering and the Geosystems Research Institute in High Performance Computing Collaboratory, Mississippi State University, Mississippi State, MS 39762 USA (e-mail: wz40@msstate.edu; du@ece.msstate.edu; fowler@ece.msstate.edu).

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.2010.2081661

II. Change Detection for Hyperspectral Imagery

The change vector analysis (CVA) of a difference image is one of the most widely used classes of change-detection methods [5]. CVA is of great use because it is not only relatively simple, straightforward, and highly implementable but also capable of providing detailed change information [6]. CVA has been widely used with multispectral data [7] and also has been employed with hyperspectral imagery [8]. Successful implementation of CVA requires elimination of changes due to atmospheric, illuminance, and environmental conditions which are irrelevant to the actual land-cover change. However, when precise atmospheric correction is not possible, radiometric normalization can function as a simplified substitute.

In radiometric normalization for hyperspectral imagery, we have two hyperspectral images, namely, reference image $\mathbf{X}$ and test image $\mathbf{Y}$, each with size $L \times N$, where L is the number of spectral bands and N is the number of pixels. Assume that each band in $\mathbf{X}$ is linearly related to all the bands in image $\mathbf{Y}$ if there is no land-cover change. In this case, we have

$$\mathbf{Y} = \mathbf{T}_{\mathbf{YX}}\mathbf{X} \tag{1}$$

where $\mathbf{T}_{\mathbf{YX}}$ is an $L \times L$ transform matrix [8] with least squares estimate

$$\mathbf{T}_{\mathbf{YX}} = \mathbf{Y}\mathbf{X}^{\mathrm{T}}(\mathbf{X}\mathbf{X}^{\mathrm{T}})^{-1}. \tag{2}$$

With a constant offset $\mathbf{d}_{\mathbf{YX}}$ for each band, (1) becomes

$$\mathbf{Y} = [\,\mathbf{T}_{\mathbf{YX}} \quad \mathbf{d}_{\mathbf{YX}}\,]\begin{bmatrix}\mathbf{X}\\ \mathbf{1}^{\mathrm{T}}\end{bmatrix} = \widetilde{\mathbf{T}}_{\mathbf{YX}}\widetilde{\mathbf{X}} \tag{3}$$

where $\mathbf{d}_{\mathbf{YX}}$ is of size $L \times 1$, $\mathbf{1}$ is an $N \times 1$ vector of all ones, and $\widetilde{\mathbf{T}}_{\mathbf{YX}}$ is an $L \times (L+1)$ transform matrix. In this affine

case, $\widetilde{\mathbf{T}}_{\mathbf{YX}}$ is estimated as

$$\widetilde{\mathbf{T}}_{\mathbf{YX}} = \mathbf{Y}\widetilde{\mathbf{X}}^{\mathrm{T}}(\widetilde{\mathbf{X}}\widetilde{\mathbf{X}}^{\mathrm{T}})^{-1}. \tag{4}$$

$\widetilde{\mathbf{T}}_{\mathbf{YX}}$ constitutes a radiometric normalization between $\mathbf{X}$ and $\mathbf{Y}$, which is assumed to compensate for changing atmospheric and illuminance conditions. Mathematically, it provides a linear regression (LR); consequently, the LR residual

$$\boldsymbol{\delta} = \mathbf{Y} - \widetilde{\mathbf{T}}_{\mathbf{YX}}\widetilde{\mathbf{X}} = \mathbf{Y} - (\mathbf{T}_{\mathbf{YX}}\mathbf{X} + \mathbf{d}_{\mathbf{YX}}\mathbf{1}^{\mathrm{T}}) \tag{5}$$

is due to any land-cover change.

While (5) can be applied across the entire image, [9] observed significant performance gain when images $\mathbf{X}$ and $\mathbf{Y}$ are partitioned into multiple, smaller subsets, or "segments," using some segmentation method with $\widetilde{\mathbf{T}}_{\mathbf{YX}}$ being designed for each segment independently. However, when areas of change are large, it may be difficult to determine the correspondence of the segments between the reference and test images. For compression purposes, the resulting irregular segment boundaries (locations) will significantly increase transmission overheads. Thus, in this research we uniformly partition a hyperspectral image spatially into a $K \times K$ arrangement of rectangular blocks of the same spatial size with each segment containing all the spectral bands, and apply (4) and (5) within each segment.

To produce a final map of changed pixels, the difference image $\boldsymbol{\delta}$ is subjected to a detailed change analysis, the goal of which being to identify the set of pixels that are "significantly different" between the reference and test images; these pixels then comprise a change mask. The change mask may result from a combination of underlying factors, including appearance or disappearance of objects, motion of objects relative to the background, or object-shape changes [10]. Ideally, the change mask should exclude "insignificantly different" pixels, such as those induced by seasonal changes. For changes that cannot be removed by radiometric normalization (e.g., seasonal changes), classification and anomaly detection are often applied to the difference image [10]. In particular, Du *et al.* [8] proposed kurtosis maximization for change analysis, detecting pixels with large kurtosis; such pixels are then considered to be "significantly different." In the remainder of this letter, we use this kurtosis-maximization approach to produce the final change mask from $\boldsymbol{\delta}$.

III. Algorithms for Multitemporal Compression

A. Temporal Decorrelation

In this section, we explore several strategies for the compression of temporal sequences of hyperspectral imagery volumes; we note that we assume that all image volumes have been coregistered in a preprocessing step. One conceptually straightforward approach would be to treat the temporal direction as yet another dimensionality to the data and deploy a suitable 4-D compression algorithm, e.g., [11]. Such a 4-D coder would generally be an extension of an existing 3-D algorithm to the next higher dimensionality, including a 4-D transform to effectuate decorrelation spatially, spectrally, as well as temporally. For hyperspectral imagery, which may be acquired on computationally constrained platforms such as airborne and satellite-borne devices, such 4-D coders may be infeasible from computational and memory-use standpoints. Instead, we assume that an encoder can store and use only one prior hyperspectral volume (the reference image) to aid in the coding of the current image (the test image). With only two images in the temporal direction, a full 4-D coder like in [11] is inapplicable since there are not enough data in the temporal direction to support a 4-D transform; consequently, we consider modifications to the 3-D coding of hyperspectral volumes to exploit the availability of the reference image.

One approach is to have the encoder code the LR difference image $\boldsymbol{\delta}$ from (5) instead of coding the current test image directly. The decoder, having access to the same reference image, can make the same prediction as the encoder; this prediction is added to the coded difference image to result in a final reconstruction of the current test image. We focus on the incremental performance of this approach, assuming that the cost in terms of bit rate of sending a high-quality reference image to the decoder is amortized over many uses of the same reference image to code multiple test images.

Prediction, as exemplified by the LR approach mentioned previously, is one common approach to decorrelation. An alternative decorrelation strategy is to deploy a transform. Having already ruled out the feasibility of a 4-D transform, we instead consider extending the spectral transform employed in the 3-D hyperspectral coder to the task of temporal decorrelation. Specifically, we form an extended 3-D volume by concatenating the reference and test images spectrally, i.e., if both images have N pixels of L bands, then the resulting concatenated image has N pixels of $2L$ bands. The spectral transform (e.g., DWT, PCA, and SubPCA) of the subsequent 3-D coder then decorrelates the data both spectrally and temporally. Upon decoding, the L bands of the current test image are extracted from the reconstructed concatenated image. An advantage of this "spectral-concatenation" (SC) approach is that, while the encoder must store and use the reference image, the decoder does not. On the other hand, SC is redundant in the sense that, if multiple test images are coded using the same reference image (as assumed before for the LR approach), the reference image is transmitted in effect multiple times to the decoder only to be discarded during reconstruction. However, the following experimental results demonstrate that this inherent redundancy does not hinder the competitiveness of the SC approach. We note that SC has been used extensively with PCA for change detection in multispectral applications, e.g., [12]–[15].

Finally, the most straightforward strategy is to have the encoder simply ignore the reference image when coding the current test image. We refer to this as the "independent" (IND) approach since the current test image is coded independently of the reference image in terms of temporal decorrelation. As can be expected, the following experimental results demonstrate that such IND coding is not as effective as the more sophisticated LR or SC approaches.

B. Change Detection and Removal

In a multitemporal set of data, changes occurring between the images are of paramount application importance; thus, it

is crucial that compression algorithms preserve such temporal changes with high fidelity. Yet, temporal decorrelation is most effective when data change little and tend to have difficulty accommodating abrupt changes. As a consequence, we propose to have the encoder remove salient changes in the current test image and replace them with values interpolated from spatially neighboring pixel vectors. The pixel vectors for the changes, as well as their spatial location in the current test image, are losslessly transmitted in the header of the compressed bitstream.

This "change-removal" (CR) approach to preserving changes in multitemporal compression can be deployed in any of the three (IND, LR, and SC) temporal decorrelation strategies discussed in Section III-A. Specifically, for the LR and SC approaches, changes are removed from both the reference and test images, the changes are replaced in both images by interpolation from spatial neighbors, and, finally, the interpolated reference image is used in the LR- or SC-based coding of the current test image. Note that the compressed bitstream stores only the changes removed from the test image in each case.

We note that the CR approach we consider here is similar to the anomaly removal proposed in [16] and [17] for the preservation of anomalies in JPEG2000 compression of hyperspectral imagery. In particular, the interpolation process we use is identical to that of [17].

IV. Experimental Results

We consider the incremental coding performance of the current test image, given that the encoder and decoder both have access to a reference image. Throughout the following experimental results, we employ JPEG2000 using several different spectral transforms (PCA, SubPCA, and DWT) along with the IND, LR, and SC strategies for temporal decorrelation, as discussed in Section III-A. For a current test image of N pixels and L bands, the JPEG2000 encoder generates a bitstream of B bits, yielding a bit rate of $R = (B + O)/N/L$ bits per pixel per band (bpppb); this is the case even for the SC-based schemes which are in effect code $2L$ bands. Here, O represents the bits needed for the multiple overheads (i.e., transform matrix, data mean, change spectra and locations, and LR coefficients for the LR case). Reconstruction quality is measured with a signal-to-noise ratio (SNR) in decibels, defined as the log ratio between the variance of the original data and the variance of the error. In the following, we present results for data from both CASI and Hyperspec sensors.

A. CASI Data

Fig. 1 shows a small field at the North Research Farm of Mississippi State University, as acquired by a CASI sensor. Seventy-two bands with high SNR were used with 2-m spatial resolution, and five data sets were acquired in consecutive months in 2002. As can be observed from the pseudo color-infrared (CIR) images in Fig. 1, the August and September images are highly similar due to their both being acquired during harvest season.

Table I presents the SNR performance at a fixed bit rate for various test images using the previous month's image as reference. We note that the SC temporal-decorrelation strategy provided the best performance for the August–September pair, while the no temporal decorrelation (i.e., IND) performs best for all the other pairs. We conclude that SC is an effective approach to temporal decorrelation, but only when the degree of change between the reference and test images is small. For more substantial amount of change, using no temporal decorrelation is more effective than either the LR or SC approaches to decorrelation. We note, too, that, in all cases, both PCA and SubPCA significantly outperform DWT for spectral decorrelation.

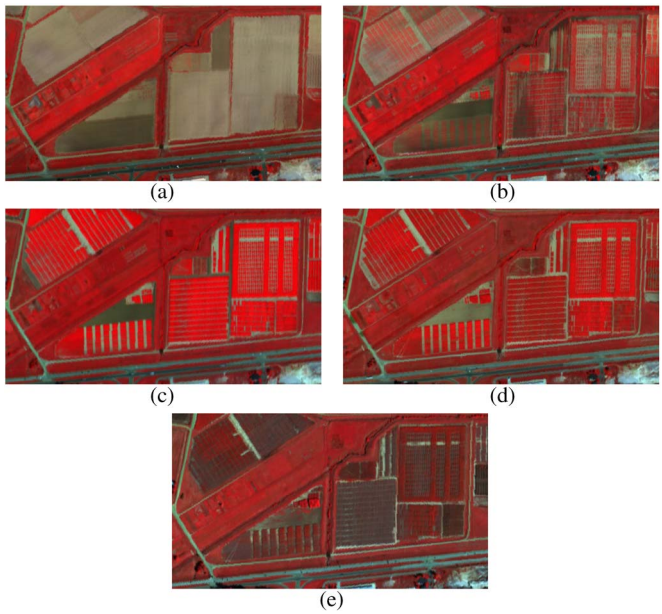

Fig. 1. CASI images taken in different months. (a) May 27, 2002. (b) July 1, 2002. (c) August 3, 2002. (d) September 10, 2002. (e) October 2, 2002.

TABLE I
SNR (IN DECIBELS) FOR THE CASI DATA AT 1.0 bpppb

Algorithms		*Reference Image / Coded Image* May / Jul	Jul / Aug	Aug / Sep	Sep / Oct
PCA	IND	**36.9**	**41.7**	38.9	**35.0**
	SC	36.2	41.5	**39.3**	33.7
	LR	36.9	41.4	38.8	34.8
SubPCA	IND	**37.0**	**41.7**	38.9	**35.0**
	SC	35.8	41.3	**39.4**	33.6
	LR	36.9	41.4	38.9	34.8
DWT	IND	**34.6**	**39.0**	36.4	**32.8**
	SC	34.1	**39.0**	**36.8**	31.4
	LR	33.9	36.6	35.9	31.2

B. Hyperspec Data

We use the multitemporal image data sets from [9], as acquired with a Hyperspec VS-25 imaging spectrograph. The first reference image [Fig. 2(a)] was taken on October 14, 2005, while the second reference image [Fig. 2(b)] was taken on September 1, 2005. The test image [Fig. 2(c)] was collected on the same day as the first reference image yet includes two tarp bundles that are absent in the reference image. All three images have 124 bands and 768 $\times$ 1024 pixels.

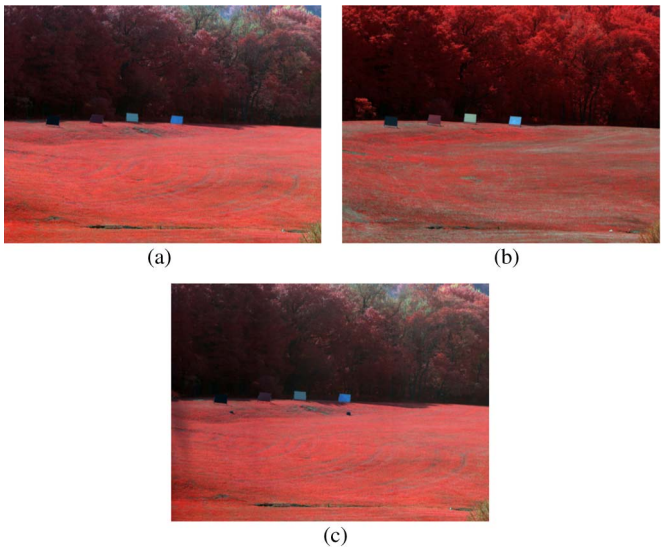

Fig. 2. Reference and test images (CIR) for the Hyperspec data set. (a) October reference image. (b) September reference image. (c) Test image.

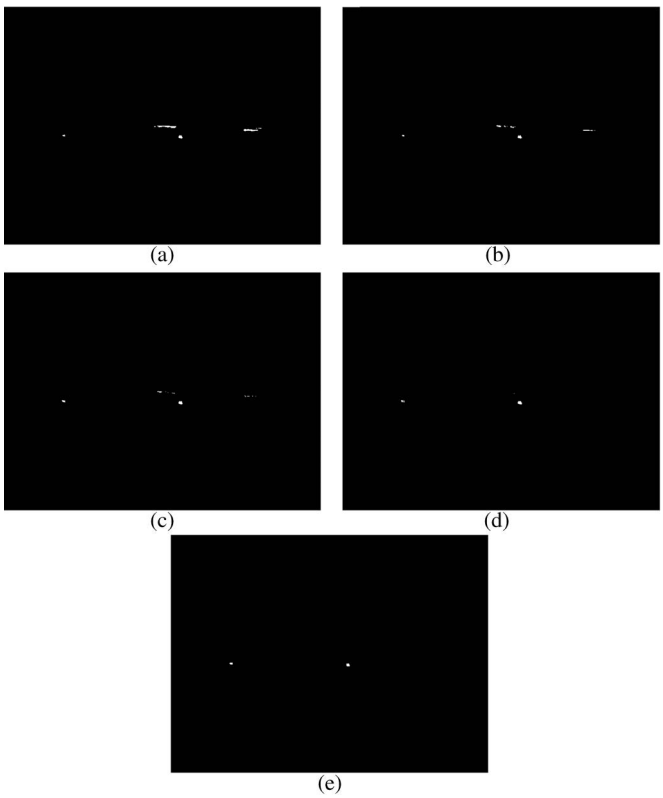

Fig. 3. Ground truth and segmented LR-based change masks. (a) $2 \times 2 = 4$ segments. (b) $4 \times 4 = 16$ segments. (c) $8 \times 8 = 64$ segments. (d) $16 \times 16 = 256$ segments. (e) Ground truth.

Fig. 3 shows the ground-truth map of the changes between Fig. 2(a) and (c), along with change masks resulting from the segmented LR-based change detection described in Section II. Specifically, the ground-truth map [Fig. 3(e)] depicts the 151 true changed pixels, while Fig. 3(a)–(d) shows the masks that result from using various segment sizes in the segmented LR change detector. It is clear that a finer spatial segmentation provides a change mask that is closer to the ground truth. However, we have observed that using too many segments

TABLE II
SNR (IN DECIBELS) FOR HYPERSPEC DATA USING THE OCTOBER REFERENCE IMAGE

		Bitrate (bpppb)					
		without CR			*with CR*		
Algorithms		0.2	0.5	1.0	0.2	0.5	1.0
PCA	IND	39.5	43.3	47.9	39.4	43.2	47.8
	SC	39.6	43.4	47.9	39.5	43.3	**47.9**
	LR	39.5	43.3	47.6	39.4	43.2	47.6
SubPCA	IND	39.5	43.3	47.9	39.4	43.3	47.8
	SC	**39.7**	**43.5**	**48.0**	**39.6**	**43.4**	**47.9**
	LR	39.5	43.3	47.6	39.4	43.2	47.6
DWT	IND	38.6	42.5	46.7	38.5	42.5	46.7
	SC	38.3	42.2	46.5	38.2	42.2	46.5
	LR	38.2	42.1	46.2	38.1	42.0	46.2

TABLE III
SNR (IN DECIBELS) FOR HYPERSPEC DATA USING THE SEPTEMBER REFERENCE IMAGE

		Bitrate (bpppb)					
		without CR			*with CR*		
Algorithms		0.2	0.5	1.0	0.2	0.5	1.0
PCA	IND	39.5	43.3	47.9	39.4	43.2	47.8
	SC	39.6	**43.6**	48.1	39.5	43.5	**48.1**
	LR	38.5	42.6	47.2	38.4	42.5	47.1
SubPCA	IND	39.5	43.3	47.9	39.4	43.3	47.8
	SC	**39.7**	**43.6**	**48.2**	**39.6**	**43.6**	**48.1**
	LR	38.6	42.7	47.2	38.5	42.6	47.1
DWT	IND	38.6	42.5	46.7	38.5	42.5	46.7
	SC	38.5	42.5	46.7	38.4	42.4	46.7
	LR	34.8	40.0	44.9	34.7	40.0	44.8

may introduce blocking artifacts which appear as lines within segment boundaries. In addition, since LR coefficients must be stored losslessly in the compressed bitstream, a large number of segments will dramatically increase overheads. As a consequence, we focus on the case of 64 segments hereafter.

Table II presents the rate-distortion performance for the compression of the test image [Fig. 2(c)] when the October image [Fig. 2(a)] was used as the reference; Table III presents the same for when the September image [Fig. 2(b)] was used as the reference. We consider the case in which changes are explicitly removed prior to compression via the CR approach outlined in Section III-B, as well as the case in which changes are left intact in the data set (i.e., "with/without CR" in Tables II and III). For CR, all 151 changed pixels, as indicated by the ground truth in Fig. 3(e), were removed. We see that, regardless of which reference image was used, SubPCA using SC temporal decorrelation achieves the highest SNR at all bit rates. Moreover, using the more distant reference image (Table III) produces 0.5–1.0-dB degradation in SNR for the LR-based temporal decorrelation. In contrast, SC performance is largely unchanged from Tables II and III, indicating that the SC approach is more effective at compensating for the irrelevant atmospheric and seasonal changes.

When changes are explicitly removed from the current test image using the CR procedure in Section III-B, the decoder has a lossless representation of the changes and their locations at its disposal in the bitstream. On the other hand, if CR is not used, the decoder would need to apply the segmented LR change detector from Section II in order to locate changes in the reconstructed data set. The performance of this change

TABLE IV
CHANGE-DETECTION PERFORMANCE AS AREA UNDER ROC CURVE FOR HYPERSPEC DATA—SEGMENTED LR-BASED CHANGE DETECTOR USED WITH 8×8 SEGMENTS

			Bitrate (bpppb)		
	Algorithms		0.2	0.5	1.0
October reference	PCA	IND	0.91	0.87	0.88
		SC	0.89	0.87	0.85
		LR	0.84	0.90	0.85
	SubPCA	IND	0.83	0.86	0.85
		SC	0.81	0.85	0.88
		LR	0.82	0.86	0.85
	DWT	IND	0.88	0.85	0.87
		SC	0.90	0.84	0.87
		LR	0.84	0.88	0.87
September reference	PCA	IND	0.65	0.90	0.81
		SC	0.85	0.89	0.86
		LR	0.80	0.84	0.75
	SubPCA	IND	0.75	0.78	0.85
		SC	0.76	0.84	0.89
		LR	0.82	0.81	0.86
	DWT	IND	0.73	0.81	0.86
		SC	0.73	0.87	0.87
		LR	0.49	0.71	0.78

detection on the reconstructed test image is presented in Table IV, wherein detection performance is measured as the area under the receiver-operating-characteristic (ROC) curve (a larger area closer to 1.0 indicates better change-detection performance). This ROC curve was estimated by normalizing a change mask to the range of [0, 1], changing the threshold gradually from 0.05 to 0.95, and then counting the resulting detection and false-alarm rates. It can be seen that the performance falls typically in the range of 0.75–0.90, somewhat short of the perfect change preservation offered by CR, which is 1.0.

V. CONCLUSION

In this letter, we have investigated multitemporal hyperspectral image compression in the case in which both the encoder and decoder use a single reference image to aid the compression of the current test image. When the data set changes little between the reference and test images, rate-distortion gain can be achieved as a result of using the reference image to remove temporal redundancy. Experimental results revealed that temporal decorrelation by means of spectrally concatenating the images, followed by spectral PCA, outperforms the commonly used linear prediction. In addition, in order to preserve temporal changes, we explicitly removed changed pixels prior to compression, storing them and their locations losslessly in the compressed bitstream. This CR technique results in only a slight decrease in rate-distortion performance while permitting perfect recovery of the change pixels at the decoder.

ACKNOWLEDGMENT

The authors would like to thank M. T. Eismann and J. Meola for providing the Hyperspec data and ground truth.

REFERENCES

[1] J. E. Fowler and J. T. Rucker, "3-D wavelet-based compression of hyperspectral imagery," in *Hyperspectral Data Exploitation: Theory and Applications*, C.-I. Chang, Ed. Hoboken, NJ: Wiley, 2007, ch. 14, pp. 379–407.

[2] B. Penna, T. Tillo, E. Magli, and G. Olmo, "Progressive 3-D coding of hyperspectral images based on JPEG 2000," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 125–129, Jan. 2006.

[3] B. Penna, T. Tillo, E. Magli, and G. Olmo, "Transform coding techniques for lossy hyperspectral data compression," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1408–1421, May 2007.

[4] Q. Du and J. E. Fowler, "Hyperspectral image compression using JPEG2000 and principal component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 2, pp. 201–205, Apr. 2007.

[5] D. Lu, P. Mausel, E. Brondízio, and E. Moran, "Change detection techniques," *Int. J. Remote Sens.*, vol. 25, no. 12, pp. 2365–2407, Jun. 2004.

[6] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307, Mar. 2005.

[7] W. A. Malila, "Change vector analysis: An approach for detecting forest changes with Landsat," in *Proc. 6th Annu. Symp. Mach. Process. Remotely Sensed Data*, West Lafayette, IN, Jun. 1980, pp. 326–335.

[8] Q. Du, N. Younan, and R. King, "Change analysis for hyperspectral imagery," in *Proc. Int. Workshop Anal. Multi-Temporal Remote Sens. Images*, Leuven, Belgium, Jul. 2007, pp. 1–4.

[9] M. T. Eismann, J. Meola, and R. C. Hardie, "Hyperspectral change detection in the presence of diurnal and seasonal variations," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 237–249, Jan. 2008.

[10] P. Coppin, I. Jonckheere, K. Nackaerts, B. Muys, and E. Lambin, "Digital change detection methods in ecosystem monitoring: A review," *Int. J. Remote Sens.*, vol. 25, no. 9, pp. 1565–1596, May 2004.

[11] G. Ziegler, H. P. A. Lensch, M. Magnor, and H.-P. Seidel, "Multi-video compression in texture space using 4-D SPIHT," in *Proc. IEEE Workshop Multimedia Signal Process.*, Siena, Italy, Sep. 2004, pp. 39–42.

[12] G. Byrne, P. Crapper, and K. Mayo, "Monitoring land-cover change by principal component analysis of multitemporal Landsat data," *Remote Sens. Environ.*, vol. 10, no. 3, pp. 175–184, Nov. 1980.

[13] J. A. Richards, "Thematic mapping from multitemporal image data using the principal components transformation," *Remote Sens. Environ.*, vol. 16, no. 1, pp. 35–46, Aug. 1984.

[14] S. E. Ingebritsen and R. J. P. Lyon, "Principal components analysis of multitemporal image pairs," *Int. J. Remote Sens.*, vol. 6, no. 5, pp. 687–696, May 1985.

[15] X. Li and A. G. O. Yeh, "Principal component analysis of stacked multitemporal images for the monitoring of rapid urban expansion in the Pearl River Delta," *Int. J. Remote Sens.*, vol. 19, no. 8, pp. 1501–1518, May 1998.

[16] B. Penna, T. Tillo, E. Magli, and G. Olmo, "Hyperspectral image compression employing a model of anomalous pixels," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 664–668, Oct. 2007.

[17] Q. Du, W. Zhu, and J. E. Fowler, "Anomaly-based JPEG2000 compression of hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 696–700, Oct. 2008.