WAVELET PRINCIPAL COMPONENT ANALYSIS AND ITS APPLICATION TO HYPERSPECTRAL IMAGES

Maya R. Gupta and Nathaniel P. Jacobson

University of Washington
Department of Electrical Engineering
Seattle WA, 98195

ABSTRACT

We investigate reducing the dimensionality of image sets by using principal component analysis on wavelet coefficients to maximize edge energy in the reduced dimension images. Large image sets, such as those produced with hyperspectral imaging, are often projected into a lower dimensionality space for image processing tasks. Spatial information is important for certain classification and detection tasks, but popular dimensionality reduction techniques do not take spatial information into account. Dimensionality reduction using principal components analysis on wavelet coefficients is investigated. Equivalences and differences to conventional principal components analysis are shown, and an efficient workflow is given. Experiments on AVIRIS images show that the wavelet energy in any given subband of the reduced dimensionality images can be increased with this method.

Index Terms— wavelet transforms, Karhunen-Loeve transforms

We explore how wavelet filtering can be paired with linear dimensionality reduction to optimally capture edge information. Experiments are presented for hyperspectral images, which contain many image bands, each of which represent reflectance information over a different spectral window. For display, the dimensionality is usually reduced to three bands which can be mapped to R, G, and B display channels. For automated classification or detection algorithms, the dimensionality is usually reduced because of the effects of the curse of dimensionality, and because image bands are often correlated.

Many hyperspectral classification methods ignore spatial information and make decisions on a pixel-by-pixel basis, including endmember unmixing [1]. However, spatial information is important for certain classification and detection tasks. Spatial context such as texture can help with accurate material identification. Higher-level classification and detection tasks benefit from incorporating spatial information [2]. We hypothesize that incorporating spatial information into the dimensionality reduction step can benefit spatially-based classification algorithms compared to using a spatially invariant dimensionality reduction.

1. WAVELET PCA

A standard method for reducing the dimensionality of hyperspectral images is principal components analysis (PCA), a data adaptive orthonormal transform whose projections have the property that for all values of N, the first N projections have the most variance possible for an N-dimensional subspace. Fig. 1 (left) shows the steps to use PCA to reduce the image dimensionality. PCA ignores spatial information; it treats the set of spectral images as an unordered set of high-dimensional pixels.

Wavelets are an efficient and practical way to represent edges and image information at multiple spatial scales. Image features at a given scale, such as houses or roads, can be directly enhanced by filtering the wavelet coefficients. For many tasks, wavelets may be a more useful image representation than pixels. Hence, we consider PCA dimensionality reduction of wavelet coefficients in order to maximize edge information in the reduced dimensionality set of images. Note that the wavelet transform will take place spatially over each image band, while the PCA transform will take place spectrally over the set of images. Thus, the two transforms operate over different domains. Still, PCA over a complete set of wavelet and approximation coefficients will result in exactly the same eigenspectra as PCA over the pixels (see the lemma).

However, PCA over a *subset* of wavelet coefficients can be used to find eigenspectra that maximize the energy of that subset of wavelet coefficients. For example, PCA on only the vertical wavelet subbands will result in eigenspectra that maximize vertical wavelet energy. More generally, we use the term *Wavelet PCA* to refer to computing principal components for a masked or modified set of wavelet coefficients to find Wavelet PCA eigenspectra, and then projecting the original image onto the Wavelet PCA eigenspectra basis. In this way, features at a particular scale are indirectly emphasized by the computed projection basis, enhancing the reduced dimensionality images without filtering artifacts.

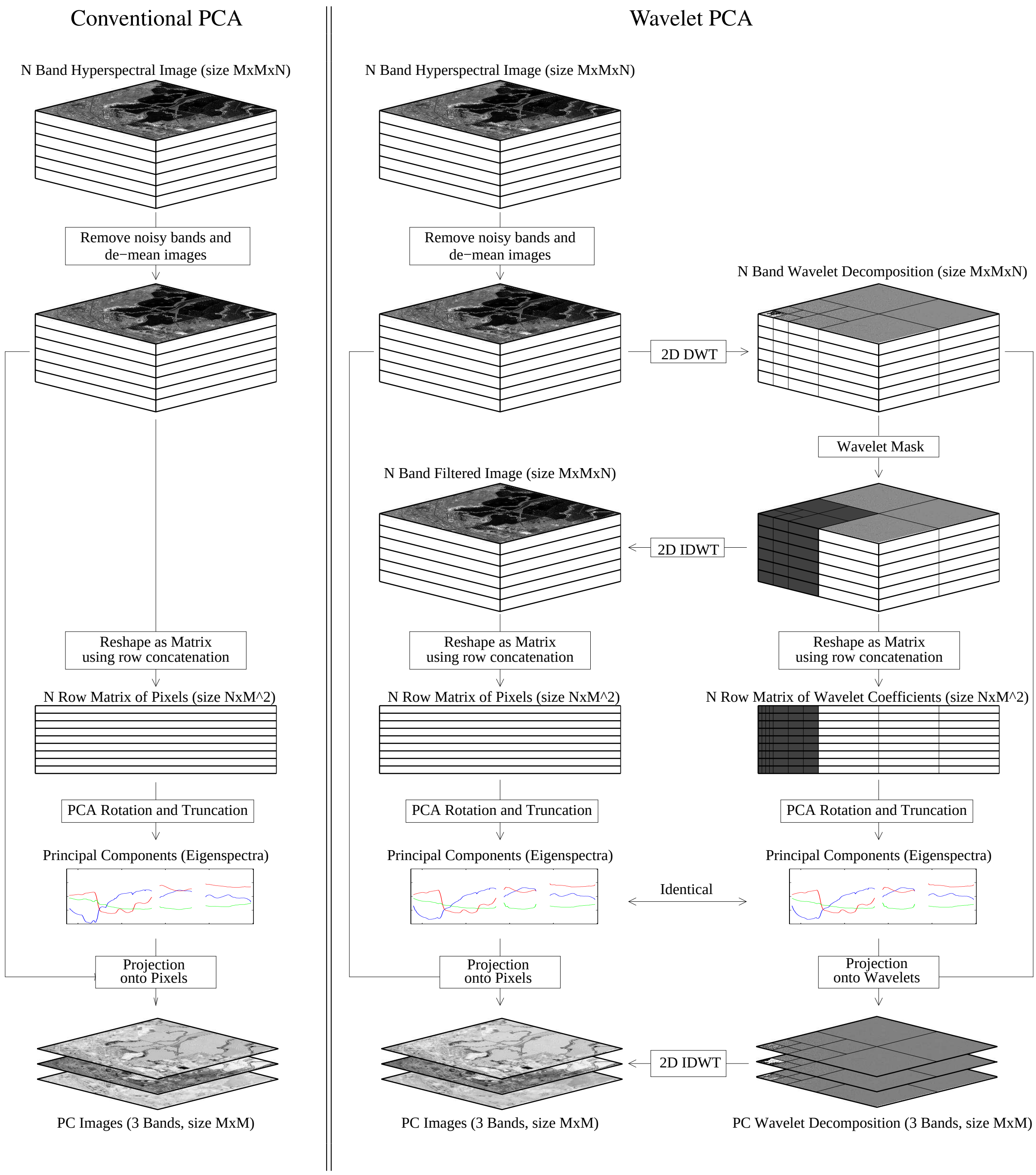

Fig. 1. **Left:** Workflow diagram for conventional PCA. **Right:** Workflow diagram for Wavelet PCA. The lemma in Section 1 establishes equivalences in the workflows.

Three different workflows are shown in Fig. 1. On the left is conventional pixel PCA. On the far right is Wavelet PCA where the eigenspectra are calculated using wavelet coefficients and the original wavelet image is projected onto the Wavelet PCA eigenspectra, and an inverse 2D DWT is performed to produce a final pixel image. In the middle, an alternate and equivalent Wavelet PCA workflow is shown, where the PCA is performed on pixels which have been wavelet filtered. Then, the original pixel image is projected onto the (identical) Wavelet PCA eigenspectra. The most efficient workflow is a fourth (equivalent) option: perform analysis in the wavelet domain and projection in the pixel domain.

Lemma: Equivalences of Wavelet PCA

With regards to the workflows shown in Fig. 1, the following are equivalent:

1. *Conventional PCA and Wavelet PCA if no wavelet coefficients have been scaled.*
2. *Wavelet PCA where the eigenspectra are computed from scaled (e.g. masked) wavelet coefficients, and Wavelet PCA where the eigenspectra are computed from the pixels resulting from the inverse wavelet transform of the modified wavelet coefficients.*
3. *Wavelet PCA where the principal component images are computed by projecting the pixel image onto the eigenspectra, and Wavelet PCA where the principal component images are computed by projecting the wavelet coefficients onto the eigenspectra then taking the inverse wavelet transform.*

Proof of the lemma is in the appendix. More generally, the proof (and lemma) hold for any orthonormal transform, not just wavelet transforms.

2. RELATED WORK

Wavelet methods have been combined with PCA in several papers. Feng et al. [3] compute eigenfaces from a mid-range wavelet subband rather than using the full image. Their approach differs from this paper in that their results are projected feature vectors with maximum variance, rather than projected images with maximum variance. Kaewpijit et al. [4, 5] develop a hyperspectral classification method that uses wavelet denoising of each pixel's spectra, followed by downsampling to reduce the dimensionality of spectra, which forms the input to a conventional PCA classifier. No spatial image information is taken into account; all operations are done on spectra, as in work by Bruce et al. [6].

3. EXPERIMENTS AND DISCUSSION

We experimented with hyperspectral imagery to explore the question, "How much difference does incorporating the spatial information into the dimensionality reduction make?" As a metric, we compare the energy in different wavelet subbands for conventional PCA and Wavelet PCA for the first three projected images. Three projections were chosen because it is common to map the first three PCA projections to red, green, and blue channels for visualization. The wavelet filtering was a binary mask that selected all the wavelet coefficients at a particular resolution and excluded wavelet coefficients that were affected by the image boundary.

Figure 2 shows a wavelet variance analysis of applying this visualization method to four different AVIRIS hyperspectral images from NASA. AVIRIS images have 224 nominal bands, but only the 190 bands which did not contain significant noise were used. In this paper, d_j, with no superscript, is the union of the horizontal, vertical and diagonal wavelet coefficients at level j. These coefficients are treated as a single subband since the image is not assumed to be purposefully oriented.

As expected, the proportion of the total wavelet variance contained in a particular wavelet subband of the first projected image is increased (compared to conventional PCA) when the eigenspectra are computed only from wavelet coefficients in that subband. However, the second and third projected images usually show a decrease in wavelet variance for the selected subband. Due to the much larger variance of the first projected image, the total variance of the first three projected images is consistently increased. This suggests that the Wavelet PCA basis captures spatial information in the first projected image that is instead captured in the second, third and subsequent projected images of conventional PCA. We hypothesize that the greater compaction of wavelet energy in the first projected images will be helpful to image processing tasks that are based on spatial information, such as detection, classification, compression, and visualization.

4. APPENDIX

Proof of Lemma: We prove the lemma for one-dimensional data; the arguments extend to higher dimensions, but at the cost of readability. Let the data matrix A be an $N \times M$ matrix representing N spectral bands for each of M pixels. Let W be an M x M wavelet transformation matrix, and let W^T denote the transpose of the matrix W. The spatial discrete wavelet transform (DWT) of A is given by $(WA^T)^T = AW^T$. Since W is orthonormal, $WW^T = W^TW = I$. Thus the covariance matrix of the wavelet coefficients is $C_W = (AW^T)(AW^T)^T$. Simplifying and using the orthonormality property, $C_W = AW^TWA^T = AA^T = C$, where C is the covariance calculated from the original pixel data. Thus, working in either domain will yield the same eigenvectors (eigenspectra) and eigenvalues, establishing the first item of the lemma.

Wavelet filtering can be expressed as matrix multiplication by a diagonal matrix because each row of the matrix A

corresponds to the signal for one spectral band. Let I_f be any diagonal matrix. The covariance matrix of the modified wavelet coefficients is

$$\begin{aligned} C_W &= \left[(AW^T)I_f\right]\left[(AW^T)I_f\right]^T \\ &= AW^T I_f (I_f)^T W A^T. \end{aligned}$$

Equivalently, performing an inverse DWT on the modified wavelet coefficients yields a covariance matrix of spatially filtered pixel coefficients:

$$\begin{aligned} C &= \left[(AW^T I_f)W\right]\left[(AW^T I_f)W\right]^T \\ &= \left[(AW^T I_f W)(W^T (I_f^T) W A^T)\right] \\ &= AW^T I_f (I_f)^T W A^T. \end{aligned}$$

The identical covariance matrices will yield identical eigenvectors, proving item 2 of the lemma. Lastly, we establish item 3 of the lemma. The ith principal component image $P_i = v_i^T A$ is equivalent to projecting the wavelet coefficients and applying the inverse DWT:

$$\left[(v_i^T AW^T)W\right] = (v_i^T A) = P_i.$$

5. REFERENCES

[1] N. Keshava, J. P. Kerekes, D. G. Manolakis, and G. A. Shaw, "Algorithm taxonomy for hyperspectral unmixing," 2000, vol. 4049, pp. 42–63, SPIE.

[2] G. Rellier, X. Descombes, F. Falzon, and J. Zerubia, "Texture feature analysis using a Gauss-Markov model in hyperspectral image classification," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 42, no. 7, July 2004.

[3] G. C. Feng, P. C. Yuen, and D. Q. Dai, "Human face recognition using PCA on wavelet subband," *Journal of Electronic Imaging*, vol. 9, no. 2, pp. 226–233, 2000.

[4] S. Kaewpijit, J. Le Moigne, and T. El-Ghazawi, "Feature reduction of hyperspectral imagery using hybrid wavelet-principal component analysis," *Optical Engineering*, vol. 43, no. 350, February 2004.

[5] S. Kaewpijit, J. Le Moigne, and T. El-Ghazawi, "Automatic reduction of hyperspectral imagery using wavelet spectral analysis," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 41, pp. 863–871, 2003.

[6] L. M. Bruce, C. H. Koger, and J. Li, "Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 40, pp. 2331–2338, 2002.

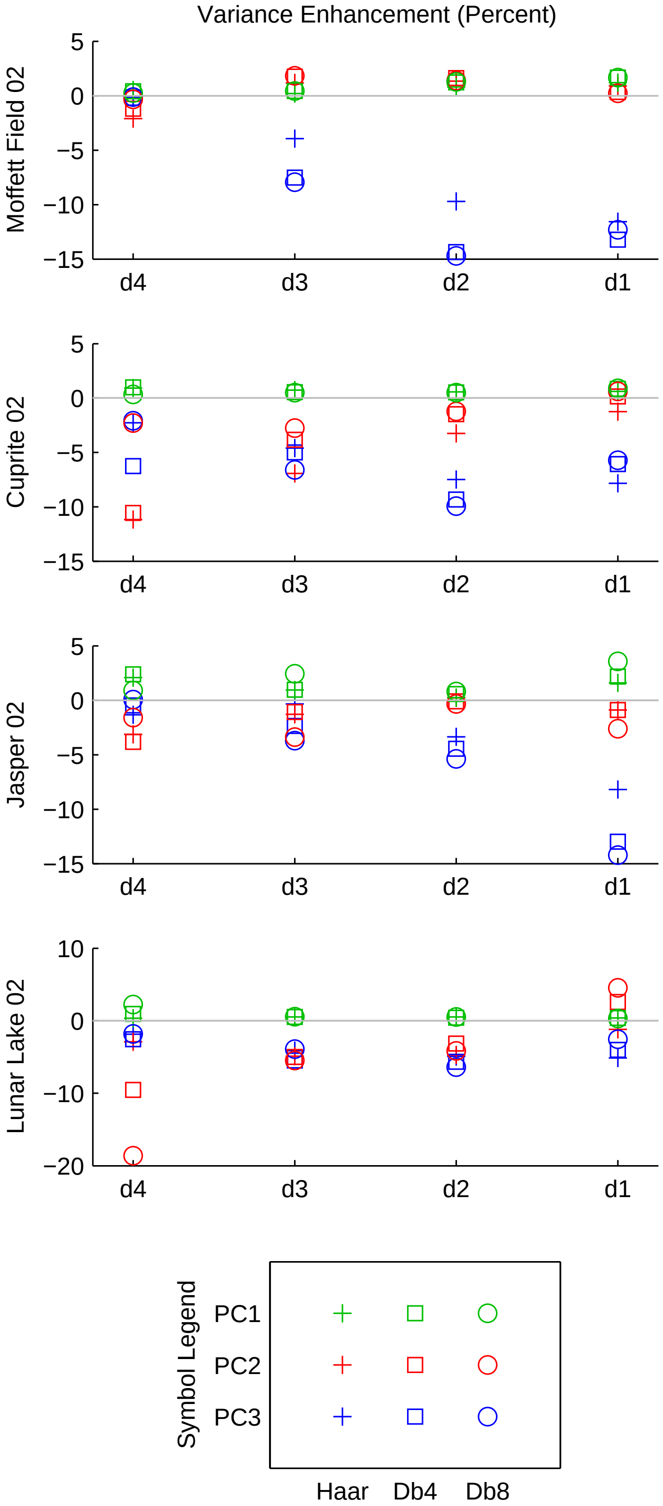

Fig. 2. Results of enhancing subband variance. Results for the first projected image are shown in green, the second in red, and the third in blue. Results for the Haar wavelet and two Daubechies wavelets, Db4 and Db8 (with 4 and 8 vanishing moments respectively) are shown. Wavelet coefficients containing boundary effects were excluded from the variance analysis.